THE UNIVERSITY
of EDINBURGH

THE UNIVERSITY *of* EDINBURGH
School of Biological Sciences

Biosensors for Heavy Metals

Jan Oltmanns

Thesis presented for the degree of Doctor of Philosophy

Edinburgh, October 2016

Declaration

I declare that this thesis was composed by myself and all the research presented in it is my own except where stated otherwise.

_____

Jan Oltmanns

Acknowledgements

This work would not have been possible without the constant help and support I received. A few words can never be enough to do this justice.

I would like to thank my supervisor Professor Christopher French for his advice, support, patience and encouragement. Thank you Chris, without you I would not be able to write these lines now.

Thank you also to the many friends I made during my time in Edinburgh. Anita, Maia, Matt, Tsveti, Pascoe, Tim, Denis, Prabu, Elvis, Joy, Maryia, Jane, Beatrice, Rocky, Pete, Chao Kuo, Julia, Alejandro, Chris M., Chris N., Marco, Giulio, Daniel, Lina, Joe and all the people back in Germany that helped me get to Edinburgh in the first place like Boris, Anna, Silke, Jens, Joe, Jogi, Leila, Ally, Sebastian, Patrice, Larissa, Sabine, Martina, Sarah, …and all the people I forgot cause I am clumsy. I have enjoyed every second with you and often would have been lost without you. A big hug to everyone! May there be many many many more happy get togethers and adventures.

I would also like to thank my partner Miguel for always having been supportive, understanding and up-lifting when I came home at the oddest of hours in the worst of moods.

I am dedicating this work to my family. My mother Gisela and father Hajo and their partners Armin and Katrin, my grandmothers Gisela and Herta, my grandfather Hans, my aunts and uncles Heike and Ferdi, Frauke and Norbert, my cousin Maike and all my other family members. I could not wish for a more supporting, understanding, generous and enabling family. I am eternally grateful for all you did and do for me and on top of all that you are fun too!

Thank you!

Table of contents

I Layman's Abstract

Heavy metals such as arsenic, lead, zinc and mercury from natural and man-made sources can be a great threat to human and animal life. They are challenging to detect, usually requiring expensive and complicated machinery. Many bacteria have developed strategies to survive in heavy metal rich environments. One of these strategies is to control the building plans (genes) for detoxification mechanisms via a two part switch, triggered by heavy metals. In this work some of these switch pairs have been employed in the design and construction of a set of biosensors aimed at the detection of heavy metals in drinking water. Biosensors usually use the ability of a component found in nature to detect whatever they are aimed at. To make this detection useful in technical applications, the detection by the natural component usually needs to be translated or transformed into a response that can be understood by electronics.

The arsenic biosensor 'AsGard' is based on one of the aforementioned switches. In the bacterium, one part of the switch releases the other part when arsenic is present, unblocking the genes when arsenic is present. For AsGard, one half of the switch was made visible by colour and fluorescence and the other part was fixed onto a plastic strip. Normally the visible part stays in the same location, bound to the fixed part. If arsenic is present, the visible part is released and washed away, thus if colour vanishes, arsenic is present. This change happens within minutes at arsenic levels corresponding to World Health Organisation guidelines for drinking water.

A second set of sensors uses the switches as they are used in the bacterium but replaces the usual genes with a reporter, meaning that if the switch is activated it will produce a machine readable optical response. These sensors have been tested in bacteria and in a more synthetic, bacteria free setup. Some of them can detect heavy metals at relevant concentrations. One of these sensors for arsenic can be adjusted to detect different levels by changing the ratio of the two switch components. Some work is shown on how to replace the reporter used by a version requiring fewer steps to be produced, possibly making future designs simpler and faster because they would require less effort by the bacterium or a synthetic system to produce the reporter once the switch is activated.

II Abstract

Heavy metals from natural and man-made sources can be a great threat to human and animal life. As small inorganic ions they are challenging to detect, usually requiring expensive and complicated machinery. Several heavy metals can accumulate in the human body, leading to long term toxic effects on the nervous system. Many bacteria have developed strategies to survive in heavy metal rich environments. One of these strategies is a bacterial operon containing genes for detoxification mechanisms controlled by a promoter and a regulatory protein. In this work some of these promoter-protein pairs, P_{ars}-ArsR, P_{copA}-CueR, $P_{merTPAD}$-MerR and P_{zntA}-ZntR from *Escherichia coli* have been employed in the design and construction of a set of biosensors aimed at the detection of heavy metals in drinking water. Biosensors usually employ biological recognition elements, transducing the signal from these to produce an output that can be integrated into electronic circuitry. The sensors presented in this work focus on reducing complexity and on providing a controlled sensor reaction.

The arsenic biosensor 'AsGard' is based on the P_{ars}-ArsR pair and functions by making the dissociation of an ArsR-mCherry fusion protein from its binding site in the P_{ars} promoter visible. In the cell, ArsR dissociates from P_{ars} upon binding of trivalent arsenic ions. Immobilising the relevant part of the P_{ars} sequence on a solid plastic support allows for the mobilisation of previously bound ArsR-mCherry proteins in the presence of arsenic to become the sensor output. The AsGard sensor detects arsenic within minutes in a concentration range overlapping with the arsenic thresholds for drinking water as set by the World Health Organisation.

Additional prototype sensors are presented bringing a reporter gene under the control of the aforementioned promoters. These sensors have been tested *in vivo* and *in vitro* in a cell free transcription translation system and partially detect metal concentrations close to relevant ranges. The P_{ars} based sensor is tuneable *in vitro* by modifying the ratio of the supplied regulatory protein ArsR and is able to detect arsenic well within the relevant range. Spinach2, a fluorescent RNA aptamer, may make future designs independent from translation, drastically reducing complexity of cell free biosensors based on cis-trans transcriptional regulation.

III Abbreviations

°C	Degree Celsius
µg	Microgram
µL	Microlitre
µm	Micromolar
A	Ampere
AA	Amino acid
AC	Affinity chromatography
ACH	ArsR-mCherry fusion protein
AGCH	ArsR-$(GGGGS)_3$linker-mCherry fusion protein
Approx.	Approximately
bp	Base pairs
bw/rev	Backward
cm	Centimetre
d	Day
$_{dd}H_2O$	Deionised water
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotidetriphosphate
EDTA	Ethylenediaminetetraacetic acid
EMSA	Electro mobility shift assay
et al.	*et alii*
fw/for	Forward
g	Gram
GFAAS	Graphite furnace atomic absorption spectroscopy
h	Hour
HF	High fidelity
His	Histidine
K	Kelvin
kb	Kilo base pairs
kDa	Kilo Dalton
KPB	Potassium phosphate buffer
KP-T	Potassium phosphate Tween buffer
L	Litre
Luc	Firefly luciferase gene
M	Molar
MCS	Multiple cloning site
mg	Milligram
min	Minute
mL	Millilitre
mm	Millimetre
mM	Millimolar
ng	Nanogram
nM	Nanomolar

OD	Optical density
ON	Over night
P	Promoter
PAGE	Polyacrylamide gel electrophoresis
PBB	Protein binding buffer
PCR	Polymerase chain reaction
pM	Picomolar
PMMA	Poly(methyl methacrylate)
PNK	Polynucleotide kinase
rcf	Relative centrifugal force
rpm	Revolutions per minute
RT	Room temperature
s	Second
sce	*sce4297* gene locus in *Sorangium cellulosum*
SDS	Sodium dodecyl sulfate
spp.	Species
t	Tonnes
TAE	Tris acetate EDTA
TBE	Tris borate EDTA
Tris	Tris-hydroxymethyl-aminohexan
U	Units
UV	Ultraviolet
V	Volt
v/v	Volume per volume
W	Watt
w/v	Weight per volume

IV List of figures

V List of tables

1 Introduction

Since the early days of science, the ability to identify the presence of certain elements and to detect and analyse chemicals has been of great interest and importance. What started as the need and desire to understand the world we live in, was lifted onto a new level of importance with the onset of the industrial revolution. Greatly increased productivity and growing abilities in chemical synthesis have allowed for stupendous leaps in how we live our lives. As the world we live in grows to be more complex, so does our need for sensing it. The detection of small, inorganic molecules – such as heavy metals – can be challenging. Detection of these molecules is traditionally based on chemical reactions or complex machinery, which can be themselves highly polluting, impractical and above all expensive (Holler et al. 1996; Nieman et al. 1998) . For these reasons, the focus of this work has been laid onto heavy metal biosensors.

1.1 Heavy metals: occurrence, toxicity and detection

There are numerous definitions for heavy metals but none has been accepted universally, making it a potentially misleading term (Duffus 2002). Most definitions imply that heavy metals are defined by their particular toxicological significance while also fulfilling some other conditions, like the formation of positive ions upon solubilisation, high density, high molar mass or high periodic table number (Sheppard 1993; Duffus 2002). For the scope of this work, only certain metals will be of relevance and for the majority of them their belonging into the vague group of heavy metals is seldom disputed (Sheppard 1993; Sarkar 2002; Bradl 2005). These heavy metals are: arsenic (As), cadmium (Cd), copper (Cu), lead (Pb), mercury (Hg), zinc (Zn) and – perhaps less clearly members of the heavy metals – gold (Au) and silver (Ag).

Heavy metals in the environment originate from natural and anthropogenic sources (Table 1.1-1 below). The ratio between these two origins is greatly dependent on the metal and the geographical location. Over the last two centuries the industrial revolution has led to dramatically increased outputs from anthropogenic sources while the growing world population and colonisation of previously uninhabited regions has led to increased exposure to natural sources (Nriagu 1996; Han et al. 2002; Callender 2003; Thevenon et al. 2011). Where water comes into contact with any metal containing minerals, pH becomes of major importance. All of the metals reviewed in this work are solubilised from rocks and mineral deposits at dramatically increased rates by acidic water (Bradl et al. 2005).

Table 1.1-1: Common sources and natural occurrence of heavy relevant for this work.
Due to some of the data referring to concentrations in solid minerals, values are given in parts per billion (ppb) and parts per million. ppb values for occurrence in water are equivalent to the same numerical value in µg/L, ppm corresponds to mg/L (Siegel 2002; Bradl et al. 2005; National Health and Medical Research Council 2011).

Metal	Natural occurrence	Anthropogenic source
Arsenic (As)	13 ppb in ocean clay and shale, 2 ppb in streams	Pesticides, electronic components, paints, wood preservatives, pesticides, herbicides, insecticides, coal combustion, mining
Cadmium (Cd)	300 ppb in shale and limestone, 0.01 ppb in streams	Batteries, pigments, metal coatings, coal combustion, plastic stabilisers, mining
Copper (Cu)	Magmatic rocks, 250 ppm in ocean clay, 94 ppm in basalt, 7 ppb in streams	Conductors, pipes, roofing, kitchenware, pigments, mining
Gold (Au)	Around 2 ppb in streams, below 100 ppb in groundwater	Mining, electronics, jewellery
Lead (Pb)	Magmatic Rocks, 30 ppm in ocean clay, 20 ppm in shale, 1 ppb in streams	Batteries, pigments, glassware, plastic, solder, pipes, mining
Mercury (Hg)	0.4 ppb in shale, 0.07 ppb in streams	Metal extraction, electronics, fungicides, mercury vapour lamps, solder, gold mining
Silver (Ag)	Up to 500 ppb in soils, 0.3 ppb in streams	Mining, electronics, tableware, chemical industry, disinfectants, preservatives
Zinc (Zn)	Magmatic rocks, 200 ppm in ocean clay, 118 ppm in basalt, 20 ppb in streams	Anti-corrosion coatings, batteries, cans, PVC, silver purification, rubber industry, paints, solder, coal combustion

Of the metals previously mentioned, copper and zinc are essential trace elements for human life (World Health Organization et al. 1996). Deficiencies in uptake of these metals can lead to adverse effects ranging from mild symptoms such as skin problems to severe conditions of anaemia and immune deficiency (Dunlap et al.

1974; Keen & Gershwin 1990). For these essential metals – which still have toxic effects when too concentrated – acceptable ranges of oral intake (AROIs) have been set, defining a range in which detrimental effects either from under- or overexposure are unlikely to appear in most humans (Sarkar 2002). For all other metals mentioned above, only recommended maximum doses have been set by several national and international agencies over the past years, indicating tolerable levels of intake and exposure for most healthy individuals (World Health Organization et al. 1996; Duruibe et al. 2007).

Since toxic heavy metals, such as arsenic, cadmium, lead, and mercury, are not metabolically degradable, they accumulate in living tissues and can lead to death or serious health issues (Sheppard 1993). Due to their solubility in water, they often disperse downstream and accumulate in clays and adsorb onto algae. This leads to their accumulation in the food chain, further increasing their impact on human and animal health (Bradl 2005).

In the following, the heavy metals relevant for this work will be discussed in regard to their occurrence in drinking water supplies, their toxicity and traditional detection methods. While a range of remediation techniques is available for heavy metal polluted solids and liquids – including technological, chemical and biological methods (Bradl & Xenidis 2005) – their discussion is beyond the scope of this work. Biosensing approaches for heavy metals are discussed in section 1.2.4, following a brief introduction into biosensors in general (1.2) and metalloregulatory proteins in bacteria (1.2.3). An overview of the limits set by various national and international agencies for heavy metals in drinking water supplies can be found in table 1.1-2.

Table 1.1-2: International limits for heavy metals in drinking water.

Where no values are given, the respective national bodies have not set guidelines. Values marked with * have been set for cosmetic reasons and not due to health implications. The approximate molar range is given for the various national and international standards. No limits have been set for Au (World Health Organization 1996; European Commission 2001; National Health and Medical Research Council 2011; Health Canada 2012; Environmental Protection Agency 2016).

Metal	WHO	Canada	EU	U.S.A.	Australia	Molar range
Ag	n.a.	50 µg/L	n.a.	n.a.	100 µg/L	0.46 – 0.92 µM
As	10 µg/L	10 µg/L	10 µg/L	10 µg/L	10 µg/L	0.13 – 0.67 µM
Cd	3 µg/L	5 µg/L	5 µg/L	5 µg/L	2 µg/L	0.018 – 0.046 µM
Cu	n.a.	1 mg/L	2 mg/L	1.3 mg/L	2 mg/L	15.7 – 31.5 µM
Hg	6 µg/L	1 µg/L	1 µg/L	2 µg/L	1 µg/L	0.005 – 0.010 µM
Pb	n.a.	10 µg/L	10 µg/L	15 µg/L	10 µg/L	0.048 – 0.072 µM
Zn	n.a.	5 mg/L*	n.a.	n.a.	3 mg/L*	45.9 – 76.5 µM

1.1.1 Arsenic

Although arsenic as an element ranks 12th in abundance in the human body, it is a potent toxin to most living organisms including humans (Eisler 1988). The greatest health risk is represented by the arsenic pollution of drinking water. This pollution occurs in the Americas and some parts of Eastern Europe, where it originates from industrial activities, but is most prominent in Asia. In India, Nepal, Thailand, Taiwan, China, Mongolia and Bangladesh a total of over 100 million people are exposed to contaminated drinking water, mainly caused by natural arsenic deposits contaminating ground waters as a consequence of insufficient geological prospecting prior to tube well drilling in developmental aid programs of the past (Cebrián et al. 1983; Smith et al. 2000; Nordblrom 2002; de Mora et al. 2011). The WHO considers drinking water with more than 10 ppb (10 µg/L or approx. 0.13 µM) of arsenic as

polluted; some older standards set this limit at 50 ppb (50 µg/L or approx. 0.67 µM, Smith et al. 2000). The continued consumption of contaminated drinking water leads to chronic arsenic poisoning, known as arsenicosis. The first symptoms occur after about 10 years and include skin lesions and colour changes in the finger- and toenails. Continued uptake leads to blood vessel diseases and cancers of the skin, bladder, kidney and lung. There are signs that various other health issues are triggered or worsened by chronic arsenic poisoning. The WHO assesses that more than 1 in 10 people will die from arsenic related cancers after continuous consumption of drinking water with 500 ppb of arsenic, approximately 1 in 100 for a contamination of 50 ppb and still about 6 in 10,000 for 10 ppb. This makes the contamination of ground water by arsenic in Bangladesh alone the largest poisoning of a population in history (Smith et al. 2000; World Health Organization 2013). As the level of arsenic contamination in drinking water can fluctuate with factors such as time of the year, amount of rain, recent flooding and changes in topography, frequent retesting becomes necessary (Berg et al. 2001; Jakariya et al. 2007). Consequently there is a need for a cheap and simple way of determining the arsenic pollution of drinking water sources by the local population. Ideally any device or method should be designed to minimise the risk of false negative results due to operating errors.

Arsenic occurs in four oxidation states of which two are bioavailable, arsenite (As(III)) and arsenate (As(V)) with As(III) being the more toxic form (Eisler 1988). Inorganic arsenic is bio-activated by undergoing a sequential reduction and methylation reactions in the body, leading to the formation of monomethylarsinic acid (MMA) and dimethylarsinic acid (DMA). MMA and DMA are genotoxic and many times more potent than arsenite at inducing DNA damage (Agency for Toxic Substances and Disease Registry 2007).

1.1.1.1 Arsenic detection

The standard method for the detection of arsenic, amongst others, in drinking water is graphite furnace atomic absorption spectroscopy (GFAAS, Rice et al. 2012, section 3500-As A2). Determination limits are usually around 5 µg/L (approx. 66.7 nM, National Health and Medical Research Council 2011). Apart from the atomic absorption spectrometer, GFAAS requires an atomization chamber with an inert

atmosphere in which the sample is vaporized off a graphite tube (furnace) by extreme heat. The tube has a limited lifetime and can need replacement after as little as 50 to 100 measurements. Since the linear measurement range is small and each instrument can behave differently due to layout and component-wear, standard curves are normally required for each measurement. GFAAS is not suitable for field use or measurement of complex samples. It requires trained personal and laboratory conditions. GFAAS is suitable for the detection of As, Cd, Cu, Pb, Ag and with limited precision for Au and Hg (Rice et al. 2012, section 3113 B). If samples are contaminated, hydride generation atomic absorption spectroscopy is an alternative with a somewhat greater sensitivity (National Health and Medical Research Council 2011; Rice et al. 2012). This method requires As(V) to be reduced to As(III), which is then converted into its volatile hydride by sodium borohydride. While eliminating influence by many contaminants in the sample through the removal of the hydrid from the sample, Ag, Au, Cu, Pb and some other metals can suppress a hydride response (Rice et al. 2012, section 3114 B).

Current field tests for arsenic pollution are mainly based on the Gutzeit method (Satterlee & Blodgett 1944; Crawford & Tavares 1974). Available test kits claim ease of use and accuracy at the relevant arsenic concentrations of 10 – 50 µg/L but large field studies showed that even field testers who were instructed in a week-long training session still produced a significant amount of false positive and false negative results (Jakariya et al. 2007). In addition, the available field test kits are too expensive for many regions in Asia (100 tests from around £ 95) and often involve production of arsine gas or mercury salts as by-products (Hach 2013). For As biosensors refer to sections 1.2.4.1.1 and 1.2.4.2.1.

1.1.2 Cadmium

Cadmium can be released into drinking water as a consequence of impurities in the zinc of galvanised pipes and solders. Industrial waste waters, contaminated fertilisers and mining present other sources. Traditionally, Cd has been used as an anticorrosive coating for steel, in batteries and electrical components and as pigment in paints and plastics (Han et al. 2002; Bradl et al. 2005; Duruibe et al. 2007).

Cadmium accumulates in the kidney and deposits are only released very slowly. The expected biological half-life in humans lies between 10 to 15 years. Long term exposure causes kidney dysfunction, bone softening and if inhaled, severe effects in the lung. The latter occurs also as a consequence of acute Cd exposure, ultimately leading to chronic obstructive airway disease, which can significantly shorten life expectancy (Fridberg et al. 1992; Adriano 2001).

Uncontaminated natural waters normally have Cd concentrations around or below 2 µg/L. Limits for drinking water have been set between 2 µg/L and 5 µg/L (approx. 18 – 44 µM) by various national and international agencies, ensuring that the life-time uptakes of the accumulating metal is kept below the critical amount of 200 mg/kg for healthy humans (Fridberg et al. 1992; World Health Organization 1996; National Health and Medical Research Council 2011; Environmental Protection Agency 2016).

1.1.2.1 Cadmium detection

The standard method for the detection of cadmium, amongst others, in drinking water is GFAAS (1.1.1.1). Determination limits are usually around 5 ug/L (approx. 44 nM, Rice et al. 2012).

Recently, nanocomposite modified carbon paste electrodes have been described, which allow determination of Cd (2.32 ug/L), Cu (9.73 ug/L), Pb (2.65 ug/L) and Hg (0.052 ug/L) at levels that are suitable for the assessment of drinking water. However, they require a platinum rod, a potentiostat/galvanostat, and purges with high purity nitrogen gas, making them unsuitable for field use (Dey et al. 2014). For a Cd biosensor refer to section 1.2.4.2.1.

1.1.3 Copper

Copper has some characteristics that make it a very popular material for a multitude of purposes. Its corrosion resistance led to its use in domestic water pipes, its high conductivity made it the most popular conductor in local electricity installations and modern electronics. Traditionally, it is used for household goods such as pots and pans and since the introduction of tap water in the home, $CuSO_4$ has been used

extensively to control algal growth in water storage units (World Health Organization et al. 1996; National Health and Medical Research Council 2011). Many minerals contain natural Cu deposits and in areas with such minerals, water from tube wells or springs can often contain Cu concentrations exceeding 5 mg/L and up to 22 mg/L, high enough to cause potential health problems (Bradl et al. 2005; National Health and Medical Research Council 2011).

While copper is a trace element crucial for human health, increased Cu uptake can have severe health effects. It is estimated that human adults require 2 – 3 mg of Cu a day. Results as to what Cu uptake has detrimental health effects differ widely. A study involving 3-month old infants showed no effects from the use of drinking water with a concentration of 2 mg/L. Other studies have shown a range of gastrointestinal symptoms like nausea, pain and vomiting due to consumed water containing 3 – 5 mg/L Cu (Pizarro et al. 1999). The use of copper kitchenware has led to severe copper poisoning including liver cirrhosis in children (Tanner 1998). Based on the known effects, a maximum tolerable daily intake for humans of 0.5 mg/kg body weight has been set by the WHO. Several national and international agencies derived guidelines for drinking water from this that are in the range of 1 to 2 mg/L (approx. 16 to 31 µM, World Health Organization et al. 1996; National Health and Medical Research Council 2011; World Health Organization 2013; Environmental Protection Agency 2016).

1.1.3.1 Copper detection

Copper can be detected by the previously mentioned GFAAS and other variants of atomic absorption spectrometry. The determination limits are well below the requirements for drinking water analysis, but the limitations due to cost and complexity persist (Rice et al. 2012, 1.1.1.1).

The previously mentioned nanocomposite modified carbon paste electrodes can be used for sufficiently precise detection of Cu (Dey et al. 2014, 1.1.2.1). For Cu biosensors refer to section 1.2.4.1.2 and 1.2.4.2.4.

1.1.4 Gold

Gold is mainly used in jewellery, dentistry, electronics, and the aerospace industry. It has also been used as a vehicle for gene delivery in gene therapy (Merchant 1998). The average abundance of Au in streams is 2 µg/L and below 100 µg/L in groundwater (Rice et al. 2012). Compounds of gold containing thiosulfate have some human toxicity and a general sensitisation to gold is possible (Merchant 1998). Due to the minimal health implications there are no limits set for Au in drinking water.

1.1.4.1 Gold detection

Gold in small concentrations (µg/L, sub-µM) is traditionally detected by various versions of atomic absorption spectrometry (Rice et al. 2012). For an example of an Au biosensor refer to section 1.2.4.1.2.

1.1.5 Lead

Lead is the most abundant heavy element in the environment (Bradl et al. 2005). Known to many as a soft metal used in roofing and solder as well as plumbing in historical buildings, it is mined and consumed in vast quantities. An estimated 3 million tons of lead have been used in 2009 for the production of batteries and accumulators, as gasoline additives, in construction, as cable coatings, in ammunition and other uses (Sheppard 1993). The release into the environment by endogenous sources is roughly equal to natural weathering from rocks (Sheppard 1993; Bradl 2005). However, atmospheric discharge by industrial processes and formerly automotive exhaust gases has led to deposition in areas far away from direct exposure through toxic discharge or natural wash-out. These effects of airborne pollution can be traced back even to Greek and Roman cultures over 2000 years ago and have reached levels close to present-day emissions repeatedly since the medieval ages (Flegal & Smith 1992; Brännvall et al. 1999).

Tap water from legacy plumbing systems, air pollution and contaminated food (through lead contaminated soil) are the most common reasons for lead exposure in the general population today (El-Nekeety et al. 2009), while for children dust contaminated by lead based paint is a major source of poisoning (Bradl et al. 2005).

Approximately 10% of ingested lead is absorbed by the body of healthy adults (around 50% in children) and most of it is deposited in the bones (Sheppard 1993). While the lead levels in blood can reduce weeks or months after lead exposure, the skeleton acts as a lead deposit and can release lead back into the blood over decades with a half-life of bone lead somewhere between 7 and 27 years (Rabinowitz 1991; Sheppard 1993).

Lead acts as an acute and chronic toxin on a cellular level, mainly, but not exclusively, affecting the nervous system (Ercal et al. 2001). Initial symptoms are psychical and include excitement, depression and irritability. Later, tiredness, joint pain, headaches and gastrointestinal symptoms occur and the effects on the nervous system intensify, slowing down nerve conduction and overall mental performance (Sheppard 1993; Patrick 2006a; Patrick 2006b; Chen et al. 2012). There is some evidence that lead can have carcinogenic effects in the kidney (Spector et al. 2011).

Lead levels in human blood are typically close to what is regarded as the threshold for toxic effects without additional dietary uptake or other additional exposure (Flegal & Smith 1992). Due to this, lead uptake should be minimised. The environmental protection agency (EPA) has set an action limit of 15 ppb (approx. 0.072 µM) for tap water, meaning that local authorities need to take measures to reduce lead levels if this threshold is crossed (Environmental Protection Agency 2016). The Drinking Water Inspectorate in the UK has set a limit of 10 ppb (approx. 0.048 µM, HKSAR Government 2015). However, the EPA states as well, that the limit of lead pollution in drinking water should be zero from a public health perspective (Environmental Protection Agency 2016). Even though the setting of acceptable lead levels above zero is due to pragmatism, these levels are still often exceeded, affecting millions of people (Olson et al. 2016). From this perspective, lead levels in drinking water should be held as low as possible.

1.1.5.1 Lead detection

The standard method for the detection of lead in drinking water is graphite furnace atomic absorption spectroscopy (GFAAS, page 6). Determination limits are usually around 5 µg/L (approx. 24.1 nM, National Health and Medical Research Council

2011). GFAAS is not suitable for field use or measurement of complex samples. It requires trained personal and laboratory conditions.

Somewhat simpler is the dithizone method. It bases on the photometric measurement of the cherry-red lead dithizonate resulting from a chemical reduction reaction and a chloroform extraction (Snyder 1947). Chloroform and other potentially harmful reagents, as well as the need for laboratory conditions and a spectrophotometer make this method unsuitable for field use. The previously mentioned nanocomposite modified carbon paste electrodes can be used for sufficiently precise detection of Pb (Dey et al. 2014; 1.2.2.1, page 8). Since lead is still widely in use in water distribution systems, easy monitoring at the point of use is the only effective way of tightly controlling lead pollution. This makes lead detection with the standard methods labour intensive and expensive, even for industrial nations. In this regard, the need for a simpler monitoring method becomes apparent. For examples on Pb biosensing refer to sections 1.2.4.2.3 and 1.2.4.2.4.

1.1.6 Mercury

While mercury can be found in many minerals and soil, the most important sources of mercury affecting human health is anthropogenic. It is mined as cinnabar, a bright to deep red ore. The ore itself has traditionally been used as a brilliant red or scarlet pigment known as vermillion. The use of mercury and cinnabar in paints has been phased out but legacy coatings persist and mercury is still used in amalgams, mirror coatings, vapour lamps, some measuring devices, pharmaceuticals, pesticides, fungicides and many other areas of industrial production and waste processing (World Health Organization et al. 1996; Rice et al. 2012). Large amounts of Hg are entering the atmosphere, streams and oceans through industrial pollution and combustion. In 1997 the United States alone estimated their annual Hg emission into the atmosphere at 158 tons (Keating et al. 1998).

Inorganic Hg^{2+} can be methylated by microorganisms in sediments when sulphides are present. This leads to the formation of methyl mercury, $(CH_3Hg)^+$, which is absorbed almost completely by the body, is very toxic and can concentrate in the aquatic food chain (Harada 1995; World Health Organization 2013). Since the 1950s there have been several cases of mercury poisoning as a consequence of

contaminated food and seafood, culminating in the mass poisonings of Minamata, Japan and the Basra poison grain disaster (Takeuchi et al. 1962; Bakir 1973; Harada 1995). In the former, industrial Hg discharge into local rivers over three decades had led to a concentration in the aquatic food chain in rivers and the local coastline so high that the local population developed a wide range of neurological symptoms from numbness to insanity, coma and death (Takeuchi et al. 1962; Harada 1995). The Basra poison grain disaster was the consequence of mercury treated seed grains being fed to livestock and used in direct food preparation. The resulting mass poisoning led to symptoms similar to those observed in Japan, affecting thousands of people (Bakir 1973).

National and international agencies have set Hg-limits for drinking water at 1 – 2 µg/L (0.005 – 0.01 µM, World Health Organization 1996; National Health and Medical Research Council 2011; Environmental Protection Agency 2016)

1.1.6.1 Mercury detection

Standard Methods for the Examination of Water and Wastewater specifies the cold-vapour atomic absorption method as the preferential method for mercury detection. Like all atomic absorption methods, the cold-vapour method requires large, costly equipment, laboratory conditions and highly trained personnel (Rice et al. 2012). The previously described nanocomposite modified carbon paste electrodes can detect Hg at sufficiently low concentrations of 0.052 µg/L at a somewhat lower technological and personal resource requirement, while still being somewhat unsuitable for field use in its current state (Dey et al. 2014; 1.1.2.1). For an example on Hg biosensing refer to section 1.2.4.2.2.

1.1.7 Silver

Silver (Ag) is widely used as a disinfectant and preservative for drinking water. Despite this, high concentrations of Ag are undesirable. One aspect is the relatively poorly understood health implication of the antimicrobial effects at very high silver levels, similar to the side effects of antibiotics. Lower levels can still lead to adverse effects if exposure takes place over an extended period. Argyria is a long known

silver induced condition in which the skin turns into a frightfully intense grey to blue (Hill & Pillsbury 1939). The EPA recommended maximum level for Ag in drinking water is 0.1 mg/L (approx.. 0.93 µM, Environmental Protection Agency 2016). Silver concentrations in natural source water are usually very low (0.2 – 0.3 µg/L). Its use as a disinfectant and pollution from industrial processes can increase these levels to 50 µg/L (approx.. 0.46 µM) and higher (Nordberg et al. 2007). Food levels are between 10 and 100 µg/kg (Gibson & Scythes 1984). It has been determined that the human lifetime no-effect level is around 10 g (Hill & Pillsbury 1939). Based on this, a daily dose of 0.4 mg (approx. 3.7 µmol) is considered safe for human consumption (National Health and Medical Research Council 2011).

1.1.7.1 Silver detection

Technical detection methods for Ag amongst other methods are graphite furnace atomic absorption spectroscopy (GFAAS, page 6) and the dithizone method (see 1.2.5.1, page 11). The detection limits are 2 µg/L (approx.. 0.02 µM) and 10 µg/L (approx. 0.09 µM) respectively (World Health Organization 1996).

The biosensor of Stoyanov et al. is able to detect Ag concentrations in the surrounding medium as low as 0.01 µM, well within the requirements for monitoring drinking water, albeit with some limitations (Stoyanov et al. 2003; 1.3.4.1.2). For an exemplary Ag biosensor refer to section 1.3.4.1.2.

1.1.8 Zinc

Zinc is included among the essential trace elements. While toxic effects due to high-dose, long-term exposure are possible, Zn deficiency is much more widespread. If toxic effects occur, they are normally a consequence of the reduced Cu uptake due to the interfering effects of a high zinc diet. In extreme cases, zinc has been linked to cell death in the brain (Plum et al. 2010).

Being relatively harmless, Zn is used in a wide range of industrial processes. Best known is the galvanising of steel for corrosion protection but Zn is also a common component in paints, cosmetics and rubber. The frequent use as a corrosion protection leads to release into potable water supplies. Unfavourable pH and other

pollutants in drinking water can lead to the deterioration of galvanised steel and iron. The resulting concentrations in drinking water can range from 2 – 11 mg/L (approx. 30-168 µM), especially if the water has been kept in galvanised water tanks. While there are no health implication limits for the Zn concentration in drinking water, aesthetic considerations have led to the introduction of limits between 3 mg/L and 5 mg/L (approx. 30 µM and 76 µM respectively). Higher concentrations can lead to unpleasant taste and appearance (World Health Organization 1996; National Health and Medical Research Council 2011).

1.1.8.1 Zinc detection

Atomic absorption spectrometry is normally chosen for the detection of zinc with a detection limit of approximately 20 µg/L (approx. 0.3 µM, Rice et al. 2012). For examples on Zn biosensing refer to sections 1.2.3.2.2 and 1.2.4.2.4.

1.2 Biosensors

Evolution has had some 4 billion years (Bell et al. 2015) to address the problem of recognising the chemical surroundings a living organism might find itself in. Microorganisms exist in almost every habitat of our planet and in order to adapt to very different conditions, are able to sense a wide range of chemicals (Horikoshi & Grant 1998; Morita et al. 1999). As such there is enormous potential in harnessing "the immensely powerful molecular recognition properties of living systems" (Budak et al. 2013). Biosensors attempt exactly this. By classical definition, "A biosensor is a device incorporating a biological sensing element either intimately connected to or integrated within a transducer. The usual aim is to produce a digital electronic signal which is proportional to the concentration of a specific chemical or set of chemicals." (Turner et al. 1987). Since the concept of a biosensor was first described by Clark and Lyons in 1962 as the combination of an enzyme with an electrochemical detector with the aim of glucose sensing in blood, the definitions have widened and today a wide range of biological elements and transducers are being employed. Generally the biological recognition elements can be classified into two classes. These are biocatalysts such as enzymes, tissue material or whole microorganisms, and bio ligands, such as trans-regulatory elements, antibodies, nucleic acids and lectins (Castillo et al. 2004; French et al. 2015). One way of categorising biosensors is by their biological recognitions element (figure 1.2-1).

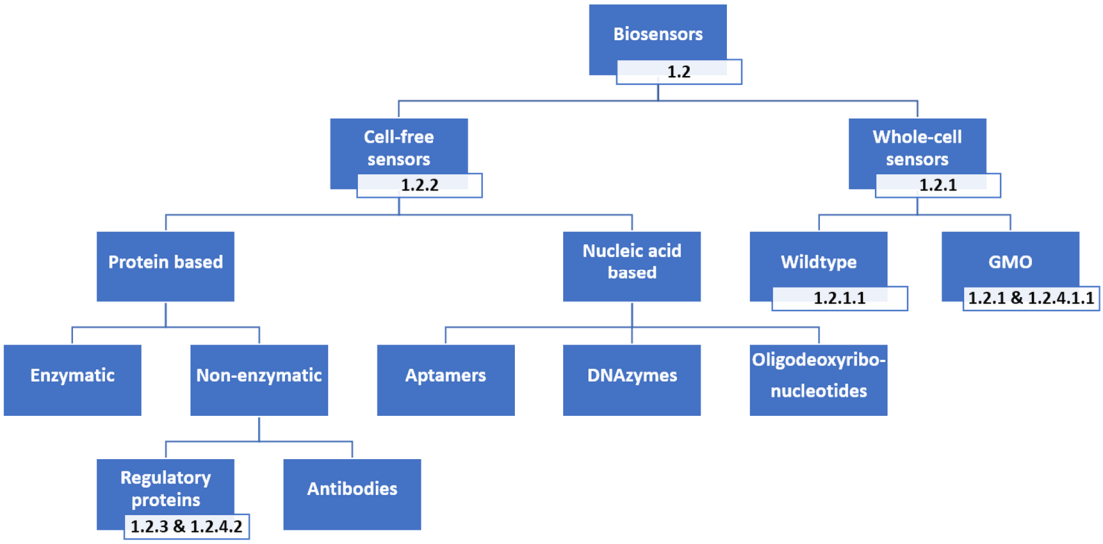

Figure 1.2-1: Classification of biosensors by common biological recognition elements (modified after Verma & Singh 2005).
Detailed subclasses of whole-cell sensors are not shown but could be divided similarly to cell-free sensors, depending on the recognition element inside the cell. Where they are discussed in greater detail, section numbers corresponding to this work are given for selected biosensor classes.

The choice of a biological element is dependent on several criteria and is tightly connected to the characteristics of the molecule(s) that are to be sensed. Molecules that cannot enter the cell, for example, cannot be sensed by a recognition system present exclusively inside the cell. Important aspects of any sensing mechanism are sensitivity and selectivity. That is, a response of sufficient strength triggered at desirable concentrations of the analyte and the specifity of this response to a single or a group of analytes. In many cases, the selectivity of biological recognition systems is greater than traditional chemical analysis. Other advantages of biosensors resulting from the biological detection element can be relatively short detections times, low manufacturing costs, miniaturisation, portability and ease of use (Yáñez-Sedeño et al. 2014; Bobade et al. 2016).
The relatively tight original definition of the transducer as an element for the generation of a digital electronic signal has been widened by approaches such as the Edinburgh Arsenic sensor (Clark & Lyons 1962; de Mora et al. 2011; 1.2.4.1.1).

These approaches are aimed at simplifying the technological components to widen the employability under adverse conditions in regard to infrastructure, financial limitations, and availability of trained personnel. Widely used transducers include electrochemical, optical, thermal and affinity monitoring (Turner 2000; Castillo et al. 2004). While the biological recognition element is crucial for the specificity, the choice of transducer has a large influence on the sensitivity of the biosensor (Turner 2000; Castillo et al. 2004)

In most scenarios the output of a biosensor is desired to be quantifiable and proportional to the amount of analyte present. The most successful biosensor, from its conception in 1962 to the present, is the blood glucose biosensor. Almost 90% of total biosensor sales are accounted for by this sensor alone (Clark & Lyons 1962; Wang 2001; Abrevaya et al. 2015). The glucose sensing system is based on the oxidation of β-D-glucose by glucose oxidase in the presence of flavin adenine dinucleotide (FAD), forming $FADH_2$. In the original system, hydrogen peroxide is produced when the $FADH_2$ reacts with oxygen. The transducer element of this sensor – a platinum anode – oxidises the hydrogen peroxide and the easily recognisable flow of electrons is proportional to the number of glucose molecules (Guilbault & Lubrano 1973; Yoo & Lee 2010). The glucose biosensor is also represents a good example for another aspect of biosensors. While many sensors are designed for continuous sensing, such as the original glucose sensor described (Clark & Lyons 1962) for blood glucose levels during surgery, some sensors are designed for a single measurement. These single measurement sensors are also referred to as “one-shot biosensors” (Turner 2000; Giedroc & Arunkumar 2007) and, looking at the glucose sensor again, are widely spread for example in the form of home use devices for the fast analysis of capillary blood samples for diabetic patients (Wang 2001).

While biosensors are often highly specific and sensitive, the incorporation of a biological component into a technical device creates problems of its own including durability and shelf-life (Rawson et al. 1989).

1.2.1 Whole-cell biosensors

Biosensors can use a wide range of biological elements and detection mechanisms associated with these. Some sensors employ entire cells, either as the sensing

element itself or to provide the necessary stable surroundings for sensing elements within the cell (Bousse 1996). While sensors have been described that are based on yeasts, algae and mosses (Ramos et al. 1993; Naessens & Tran-Minh 1998; Lehmann et al. 2000; Durrieu & Tran-Minh 2002), most whole-cell sensors employ bacterial microorganisms. Bacterial cells have the advantage that they supply a relatively stable environment to any sensing mechanisms contained in them, while being able to tolerate environmental conditions that might destroy the sensing element if it was exposed to them (Castillo et al. 2004). Since specific metabolic pathways can be used in microorganisms, the resulting microbial sensors have the potential of being more selective for compounds or pollutants that cannot be measured by simple enzyme reactions, such as aromatic compounds and heavy metals (Riedel et al. 2002; Verma & Singh 2005).

Growing bacteria is quick, easy and affordable compared to the purification necessary if only certain elements of any given cell are to be used. At the same time whole cell approaches allow for complex recognition systems involving various biological components that would otherwise be difficult to assemble. As such the advantages of whole-cell sensors lie mainly in their self-replication, self-assembly and the resulting ease of manufacture (Simpson & Sayler 2001; Yagi 2007). This advantage becomes more significant when the sensing mechanism employs multiple stages that require steps such as transcription / translation or other complex multi-step reactions, which are carried out efficiently in a cell (Castillo et al. 2004).

The most common outputs of whole cell biosensors lead to a change in the surrounding media which is then registered by technical means. Examples are pH change, colour change, light emission or other changes in the chemical composition of the extracellular or intracellular environment (Bousse 1996). Most of these signals derive from the secretion of chemical compounds or enzymes and as such require a relatively long incubation time in the range of several hours in which the necessary compounds or enzymes are produced and secreted by the bacterial cell. Conversely, to allow multi-use biosensors the produced molecules must be neutralized or inactivated when the triggering stimulus is reduced or no longer present, leading to further delays (Baker & Gough 1996; McAdams & Arkin 1997; Daunert et al. 2000).

Recently increased efforts have been made in mimicking semiconductor devices within a microbial cell in terms of processing various inputs and creating logic gates (Simpson & Sayler 2001; Bonnet et al. 2013). Moving some of the functions normally performed by silicon based semiconductors into biological components allows for a higher degree of integration (Simpson & Sayler 2001). However, the necessary molecular engineering of the cells employed in such sensors is relatively complicated. Despite the availability of fully sequenced genomes for many microorganisms, gene functions and parts of the metabolism are still not fully understood.

Recent advances in the creation of synthetic bacterial lifeforms have led to JCVI-syn3.0, a minimal genome organism consisting of only 473 genes at a total genome size of 531 kb. Yet, 149 of its genes fulfil unknown biological functions and the vast majority of the genes that were deleted from the genome of the originating organism *Mycoplasma mycoides* were of unknown function in the first place. The creators of JCVI-syn3.0 have summed this up quite clearly: "The minimal cell concept appears simple at first glance but becomes more complex upon close inspection. In addition to essential and nonessential genes, there are many quasi-essential genes, which are not absolutely critical for viability but are nevertheless required for robust growth (…), suggesting the presence of undiscovered functions that are essential for life." (Hutchison et al. 2016). The struggle in understanding and engineering bacterial cells shows that the metabolism of microorganisms is perhaps not as easy to manipulate as often claimed (Su et al. 2011) and its complexity can present an obstacle in the design of some biosensors. As such, the convenience of a whole-cell biosensor approach in regard to the "biological package" that protects, replicates and manages the sensing components (Castillo et al. 2004) also introduces a degree of uncertainty and complexity, which might not always be desirable.

The "biological package" comes with another limitation: One could say it has an expiration date printed on it that significantly shortens if the package is kept under the wrong conditions. Any biosensor design that is based on living microbial cells faces the problem of a comparatively short shelf life compared to designs that are cell free (Bjerketorp et al. 2006). As the cells replicate they require a constant supply of nutrients, trace elements, sometimes oxygen and other substances necessary for

their proliferation. At the same time waste products need to be removed from the environment the cells are living in. Failure of supply or removal usually leads to a stagnation in growth and ultimately to cell death. Due to this, sensors containing live cells require either a steady or dynamic control of the growth conditions to allow for continuous culture, making them technologically complex and prone to contamination. Continuous cultivation can further lead to genetic drift, causing the biosensor to fail, either due to a dysfunctional sensing element or an inability to transduce the signal (Pooley et al. 2004; Endo et al. 2010).

As an alternative to continuous culture, a range of approaches have been developed to supply biosensing bacteria on demand. These include air drying (Hays et al. 2005), freeze drying, vacuum drying, encapsulation in organic or inorganic polymers (Bjerketorp et al. 2006) and the use of bacterial endurance forms, such as spores (Date et al. 2010). While freeze drying and vacuum drying can drastically increase the shelf life of whole cell biosensors, some microorganisms show significantly reduced viability rates upon rehydration. Encapsulation overcomes this loss of viability but does not hinder unwanted bacterial growth to the same extend (Bjerketorp et al. 2006). Spore based approaches have been shown to be practical in terms of shelf life and viability as well as speed of recovery of vegetative cells (Date et al. 2007; Date et al. 2010). However, only a limited number of well-studied microorganisms combine the ease of genetic manipulation desired for construction of a biosensor with the ability to sporulate. Summarised, the aspect of shelf life can be challenging for whole cell biosensors, sometimes making them unsuitable for long term storage, transport and field use.

The very advantage of the cell as a packaging for the sensing mechanism can also become a disadvantage. Analytes need to be able to enter the cell via diffusion, nonspecific uptake or active transport to be available for any sensing mechanism that is based inside the cell. If membrane permeability for an analyte is not sufficient, alternative approaches include the introduction of a transport mechanism that will import the analyte in sufficient concentrations into the cell or the specific allocation of the recognition element to the outside of the cell. Both approaches hold major risks of complications, as manipulations of microbial membranes require careful engineering to prevent overcrowding or impeding native membrane functions

(Minton 2000; Katzen et al. 2009). Due to these problems, whole-cell biosensors are predominantly employed for the sensing of small molecules, which can readily enter the cell or are imported into it naturally (Bousse 1996; Daunert et al. 2000; Date et al. 2007; Struss et al. 2010; Bereza-Malcolm et al. 2015).

Perhaps one of the most eminent limitations of whole-cell biosensors is human-made:

Since they often incorporate a genetically modified organism, strict regulations apply in many countries regarding field usage (Frey 2007).

1.2.1.1 Whole-cell natural biosensors for heavy metals

There are several designs of whole-cell biosensors employing unmodified organisms; two examples focussing on heavy metal detection will be discussed here. A general advantage of unmodified whole-cell systems is the relative independence from legislative regulation, given that the employed organisms are not pathogenic or otherwise dangerous.

Yamasaki et al. have described a biosensor based on the inhibition effect of copper ions on bacterial growth. Bacteria isolated from a Portuguese lagoon were grown on a piezoelectric quartz crystal. Through monitoring the crystal's frequency, the bacterial growth could be measured in real time. Addition of copper (Cu(II)) to the growth medium above certain concentrations led to a decrease in bacterial growth. Copper concentrations as low as 18 ppm (approx.. 28 mM) could be measured this way (Yamasaki et al. 2004).

Ramos et al. have developed a microbial electrode biosensor based on the ability of mosses to accumulate metals. The moss *Sphagnum sp* was incorporated into a carbon paste electrode at a ratio of 10%. Due to the natural ability of this and other mosses to accumulate metals, lead present in the sample was concentrated at the electrode. After a preconcentration time of 15 minutes, lead concentrations as low as 2 µg/L could be measured at sufficient precision via anodic stripping differential pulse voltammetry (Ramos et al. 1993).

Microbial electrodes as described above can be susceptible to interference and contamination as they are based on relatively unspecific toxicity or accumulation processes. Sensors employing genetically engineered microorganisms (GEM) can

partially overcome this limitation by selectively producing a quantifiable signal (Verma & Singh 2005).

1.2.2 Cell-free biosensors

While cell based sensors have their greatest advantage in the self-replicating package of microbial cells (Simpson & Sayler 2001), cell-free sensors trade this convenience for a range of other advantages.

Since the sensing element is not hidden inside a living cell, analytes can access it faster and more easily, potentially leading to higher sensitivity (Richins et al. 1997; Harms et al. 2006). The absence of ongoing metabolism reduces the complexity of the surrounding medium for the recognition element. As a consequence interactions with other proteins, nucleic acids and chemicals are potentially reduced. However, the exposed detection element can increase the requirements for sample preparation in comparison to whole cell sensors. Conversely, this can lead to an overall greater complexity of use and leads to the need for systems reliably showing the unimpeded functionality of the sensing element (Ko Ferrigno et al. 2016).

Generally the production of recognition elements for a cell free sensor will include the cultivation of organisms incorporating the element and a later purification step. This increases the complexity and cost of the production process in comparison to the fully self-replicating whole-cell system (Yagi 2007; Su et al. 2011). Yet, as previously discussed, maintaining cell viability can be an issue and makes long term storage, transport and use under adverse conditions problematic (Bjerketorp et al. 2006). Cell free systems based on relatively robust detection elements can survive more extreme conditions and can require less preparation for long term storage then a living organism does (Bjerketorp et al. 2006; Kawakami et al. 2010). Of course this advantage is greatly dependent on the robustness of the employed recognition element and can vary greatly.

One aspect that can make whole-cell sensors difficult to design, is the need to engineer all parts of the sensing “circuit” to function within the cell (Simpson & Sayler 2001). When the cell replicates each new cell needs to contain the same fully functional sensor without any need for outside influences on the cell with the exception of an antibiotic selection in some cases. This means that in complex

sensors a multitude of genetic manipulations can become necessary. If certain components should be present in precisely controlled ratios to other components, such as cis/trans elements, reliable promoters need to be employed and often the cultivation conditions can still change expression ratios in unfavourable ways (Merulla, Buffi, et al. 2013; Zhang et al. 2016). Since a cell free sensor normally only employs purified components of the cell, sensor elements that might be difficult to combine in a living cell can simply be combined by mixing the independently produced components in the desired ratio. This also means that changing external conditions can only have an influence on the functionality of the sensor components, not on their actual presence, thus eliminating a factor of uncertainty.

Depending on the individual design, cell-free systems can potentially be smaller and more portable than whole-cell systems (Pardee et al. 2014), especially when considering the need for containment in any whole-cell sensor that contains genetically modifies organisms (GMOs). As such one of the biggest advantages of a cell-free biosensor in regards to field use is its exemption from the tight legislative regulation of GMO release (European Commission 2001; Pellinen et al. 2004).

1.2.3 Non-enzymatic metalloregulatory protein families

Since many of the biosensors described in the following parts of this introduction are bases on metalloregulatory proteins, a short discussion of their function is given here. Many bacteria have developed systems to survive in environments rich in heavy metals (Kaur & Rosen 1992; Nies 2003). Most of these systems function via bacterial operons, which are controlled by metal inducible promoters. The genes in these operons often code for detoxification systems based on metal transporting ATPases and regulatory proteins acting as metal responsive trans factors, controlling the expression of the operon in a feedback loop (Nies 2003; Moore & Helmann 2005). Other gene functions include sequestration, storage and intracellular trafficking (Osman & Cavet 2010). The metalloregulatory trans factors show high affinity metal-binding properties on one hand and specific interactions with DNA on the other hand (Chen & Rosen 2014). This potentially allows bacteria precise control of the expression of the genes under control of these trans factors (Busenlehner et al. 2003).

Several families of soluble metalloregulatory trans factors are known in bacteria. These include ArsR-SmtB, MerR, CsoR-RcnR, CopY, DtxR, Fur and NikR While the members of the same family can sense different metals, a specific metal can also be sensed by members of different families (Osman & Cavet 2010). How this metal specifity is realised can differ between the families and is not fully understood for some of them (Pohl et al. 2003; Osman & Cavet 2010). This work focuses on trans factors from the ArsR-SmtB and MerR families. Fortunately, the metal specifity is relatively well understood in these protein families (Osman & Cavet 2010).

1.2.3.1 The SmtB-ArsR family of metalloregulatory proteins

A tightly-controlled system for arsenic detoxification was first characterised in *E. coli* (Kaur & Rosen 1992). Arsenic enters the bacterial cells through nonspecific transport via the active phosphate import system (Eisler 1988). Arsenic resistance is based on the specific export of arsenite (As(III)) and arsenate (As(V)) (Silver et al. 1981; Rosen & Borbolla 1984; Nies 2003, refer to section 1.2.1, page 5 for details regarding the toxicity and chemistry of Arsenic). The mechanism is best-studied in *E. coli* but is widely similar in most microorganisms (Carlin et al. 1995; Moore & Helmann 2005). The *E. coli ars* operon contains five genes – *arsRDABC* – under the control of the promoter P_{ars} (figure 1.2-2). The genes *arsR* and *arsD* encode regulatory proteins while *arsA* and *arsB* encode for the two subunits of the arsenite pump (Francisco & Hope 1990; J Wu & Rosen 1993; Busenlehner et al. 2003; Nies 2003). An arsenate reductase encoded by *arsC* reduces As(V) to As(III), which can be transported by the arsenite pump (Gladysheva et al. 1994).

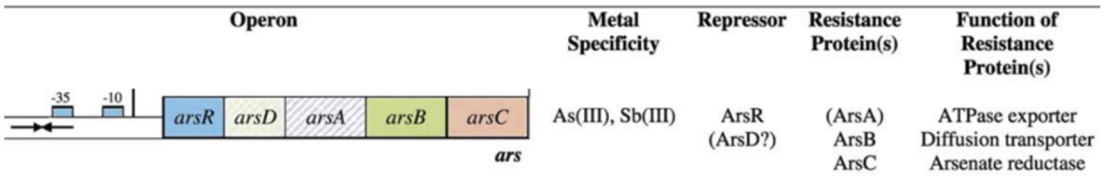

Figure 1.2-2: *ars* operon regulated by ArsR and general operon structures which confer metal resistance in bacteria. ArsD and ArsA given in bracket as not all operons contain the coding genes. (reprinted modified with permission from Busenlehner et al. 2003, copyright 2003 Federation of European Microbiological Societies)

The regulatory protein encoded by *arsR* is a transcriptional regulator of the ArsR subgroup of arsenic and antimony binding transcription factors belonging to the SmtB-ArsR family. ArsR is active as a homo-dimer. Binding of As(III) to ArsR reduces the binding affinity of the ArsR dimer to its specific binding site downstream of the -10 region in the promoter-operator (P_{ars}) of the *ars* operon through an allosteric conformational change. The dissociation of ArsR from the DNA allows for transcription of the *ars* operon to take place (Busenlehner et al. 2003).

The binding site of ArsR in the P_{ars} sequence is an imperfect 12-2-12 inverted repeat (figure 1.2-2; Xu et al. 1996).

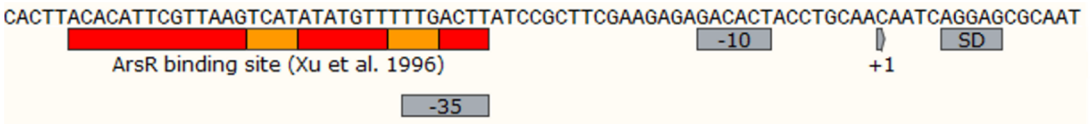

Figure 1.2-3: Regulatory region of the chromosomal *ars* operon in *E. coli*.
Marked in red is the binding site for ArsR as defined by DNase I foot printing. Within this site marked in orange are the contact points between ArsR and DNA. With grey boxes, +1 indicates the transcription start site, -10 and -35 the corresponding promoter elements and SD the most likely Shine-Dalgarno sequence (modified after Xu et al. 1996). The sequence has been shortened in front of the ArsR binding site for enhanced visibility. The unshortened total sequence length is 118 bp.

ArsR is a small cytosolic protein consisting of 117 amino acids (*E. coli*) in a conserved α1-α2-α3-αR-β1-β2-α5 structure. It binds DNA through a helix-turn-helix motif (α3-turn-αR) and specific residues on αR. No crystal structure has been published for ArsR, but the structure of highly similar SmtB has been resolved and is shown in figure 1.2-4.

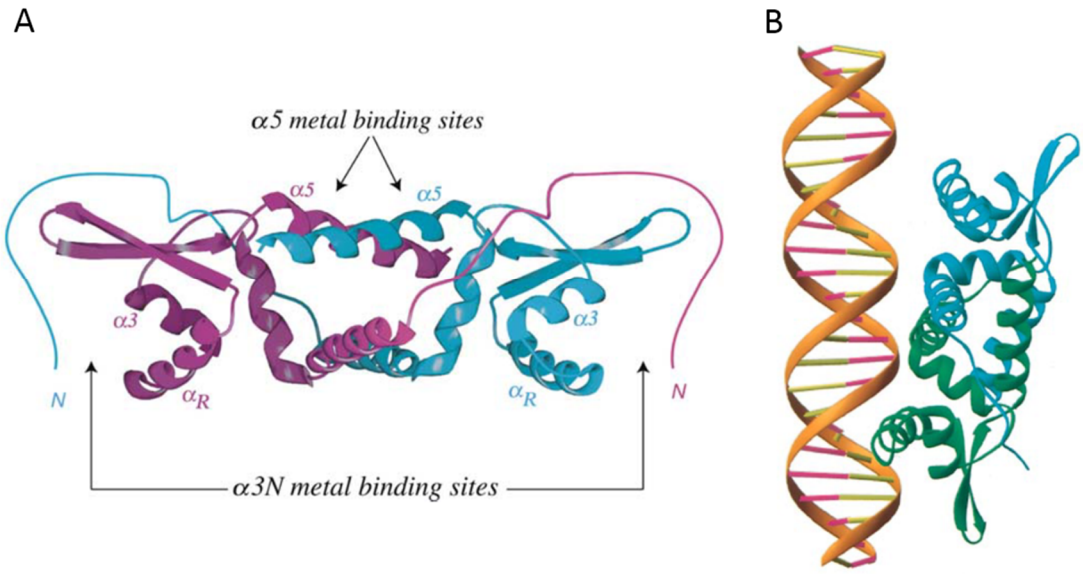

Figure 1.2-4: Crystal structure of the ArsR homologue SmtB from *Synechococcus* PCC7942 and model of SmtB binding to DNA. A) Ribbon representation of the crystal structure of the apo-SmtB homodimer with one monomer shaded purple and the other blue. Note that ArsR does not possess an N-terminal tail or α5 metal binding sites and instead has 21 additional AAs at the C-terminal sequence end (reprinted modified with permission from Busenlehner et al. 2003, copyright 2003 Federation of European Microbiological Societies). B) Model of SmtB binding to DNA based on the similarity to the HNF-3/*forkhead* DNA recognition motif and its resolved crystal structure (Clark et al. 1993). Monomers of the SmtB homodimer are shaded blue and green and DNA shown is an idealised β-DNA model (reprinted modified with permission from Cook et al. 1998, copyright 1998 Academic Press Limited).

This binding mechanism is very similar in structure to other bacterial transcriptional regulators like the catabolite activator protein (CAP) and the mercury responsive regulator protein MerR. Metal binding takes place at an ^{30}ELCVCD^{35} motif, the so called “metal binding box”, in the α3 helix (Shi et al. 1994; Busenlehner et al. 2003; figure 1.2-3).

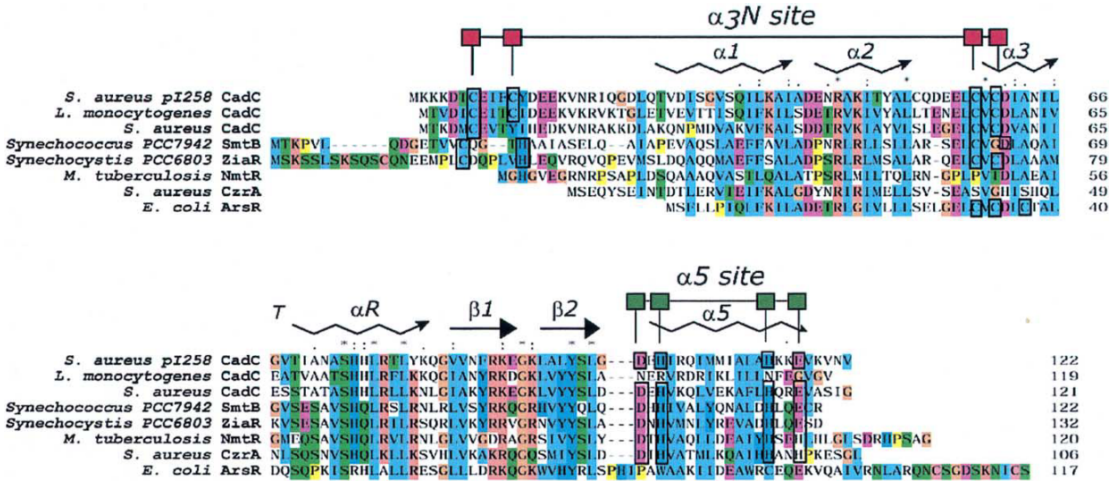

Figure 1.2-5: Proposed α3N and α5 metal binding sites for functionally characterized SmtB/ArsR proteins.

A multiple sequence alignment of several functionally characterized SmtB/ArsR proteins generated by ClustalW (Swiss-Prot, TrEMBL, or GenBank accession numbers are in parentheses): *S. aureus* pI258 CadC (P20047), *L. monocytogenes* CadC (Q56405), *S. aureus* CadC (P37374), *Synechococcus* PCC 7942 SmtB (P30340), *Synechocystis* PCC 6803 ZiaR (Q55940), *M. tuberculosis* NmtR (NP_218261.1), *S. aureus* CzrA (O85142), and *E. coli* R773 ArsR (P15905). The proposed α3N and α5 metal binding sites are denoted with the secondary structure of apo-SmtB indicated above the alignment, with the conserved ELCV(C/G)D metal binding box located in the α3 helix. Conserved metal binding residues are boxed and the amino acids are colour-coded as follows: orange-basic; purple-acidic; green-polar neutral; blue-ionisable aromatic; light blue-non-polar aliphatic; yellow-proline; brown-glycine; black-divergent. In both α3N and α5 sites, the N-terminal ligand pair is known or predicted to be derived from one monomer, with the C-terminal ligand pair derived from the other monomer within the homodimeric protein to create an inter-subunit metal binding site. The α3N metal binding ligands are indicated with red boxes in the secondary sequence designation, while the α5 ligands are indicated with green boxes (reprinted with permission from Busenlehner et al. 2003, copyright 2003 European Microbiological Societies).

Replacing the two cysteines in the ELCVCD motif with non-metal-liganding residues leads to constitutive binding of ArsR to P_{ars}, regardless of the presence of arsenite salts. A third cysteine on the α3 (C^{37}) is involved in the arsenic binding process, forming a trigonal metalloid binging site, but is not considered to be essential for the conformational change upon As(III) binding that leads to the release

of ArsR from P_{ars} (Shi et al. 1996; Osman & Cavet 2010) . Affinity of the ArsR dimer to P_{ars} when bound to As(III) is decreased (table 1.2-1).

Table 1.2-1: Thermodynamic parameters of the ArsR protein (Merulla et al. 2013).

Constant	Description	Value
$\mathbf{K_A}$	Equilibrium constant for ArsR binding to its DNA binding site	$2x10^{12}(M^{-1})$
$\mathbf{K_C}$	Equilibrium constant for binding of As(III) to ArsR	$2x10^{14}(M^{-2})$
$\mathbf{K_D}$	Equilibrium constant for ArsR-As(III) binding to its DNA binding site	$2x10^{9}(M^{-1})$

Since other members of the SmtB-ArsR family sense metals in ways that can be slightly different to ArsR, a "theme and variations" model has been proposed (Busenlehner et al. 2003). It defines the α3 site (ELCV(C/G)D or metal binding box) or α3N site (if the metal binding box interacts with residues on an N-terminal arm) as well as an α5 site with its own binding motif. One or both of these sites can be relevant for metal binding in members of the SmtB-ArsR family. *E. coli* ArsR, which is the ArsR employed in this work, does not have an N-terminal arm and the α5 site is absent, thus only the α3 site is relevant for metal binding (Busenlehner et al. 2003). Metals that are bound by *E. coli* ArsR are trivalent arsenic (As(III)) and antimony (Sb(III)). The specifity for these metals is conferred mainly by their thiophilic characteristics, making them ideal binding partners for sulfur ligands like the cysteines in the metal binding box. The spatial orientations of the box and ligand increase this specificity (Ruan et al. 2006; Chen & Rosen 2014).

1.2.3.2 The MerR family of metalloregulatory proteins

The MerR family of transcriptional regulators consists of proteins that are defined by their sequence similarity in the first 100-110 N-terminal amino acids. This region forms a helix-turn-helix motif (figure 1.2-6, page 31), specifically binding regulatory regions in promoter DNA consisting of imperfect inverted repeats (Stoyanov et al. 2001; Brown et al. 2003). In contrast to SmtB-ArsR proteins, the regulators within

the MerR family act as activators by realigning the operator DNA into a more ideal conformation for the initiation of transcription through a characteristic DNA distortion mechanism. This is possible because the controlled promoters are suboptimal in that the spacing between the -35 and -10 elements is longer (19-20 bp, figure 1.2-6, page 32 and figure 1.2-8, page 34) than the optimal 17±1 bp, placing these elements on opposite faces of the DNA double helix and making recognition by the σ^{70}-factor inefficient. MerR family proteins, when bound to their effectors, partially unwind this region on the promoter DNA, bringing the -35 and -10 elements onto the same face of the DNA molecule and into greater proximity, thus enabling more efficient transcription of the controlled genes (Frantz & O'Halloran 1990; Newberry & Brennan 2004; Philips et al. 2015). While most regulatory proteins known today bind or unbind DNA to confer activation or repression functionality, MerR family members can stay bound to the palindromic binding site in the promoter region regardless of effector binding. Recent research indicates that they have an additional repression activity through further bending the DNA into an unfavourable conformation for polymerase binding when no effector is bound (Philips et al. 2015).

MerR family proteins form head-to-tail homodimers via a coiled-coil formed by a central helical domain. This brings the DNA binding domain at the N-terminal end and the sensing region at the C-terminal end into proximity. Intimate contacts between these regions may play an important role in the DNA rearranging functionality of MerR family proteins (Heldwein & Brennan 2001; Watanabe et al. 2008; Kumaraswami et al. 2010; Philips et al. 2015; figure 1.2-9, page 35).

Members of the MerR family are known to control stress responses to a diverse range of effectors from oxidative stress to exposure to antibiotics and metal ions. This leads to great variation in their C-terminal sensing regions (Couñago et al. 2016). This work focuses on the metal-responsive members of the MerR family. These proteins share a high C-terminal sequence identity while selectively recognising different metals (Brown et al. 2003). The metal binding specifity of these MerR family proteins is not fully understood but is believed to be due to metal-coordinating residues at the C-terminal domain. Mutational studies have shown that the correct positioning and orientation of these residues is critical for specific metal

binding and regulator activity (Hobman et al. 2005; Philips et al. 2015). The structure of MerR family metalloregulatory proteins is shown in figure 1.2-9 on page 35 using the example of CueR and in an amino acid alignment in figure 1.2-4 below. The alignment shows the great sequence identity in the DNA binding Helix-turn-Helix motif and the overall sequence similarity between the three metalloregulatory proteins CueR, ZntR and MerR, including the conserved cysteine residues thought to be crucial for metal binding (Outten et al. 2000; Philips et al. 2015).

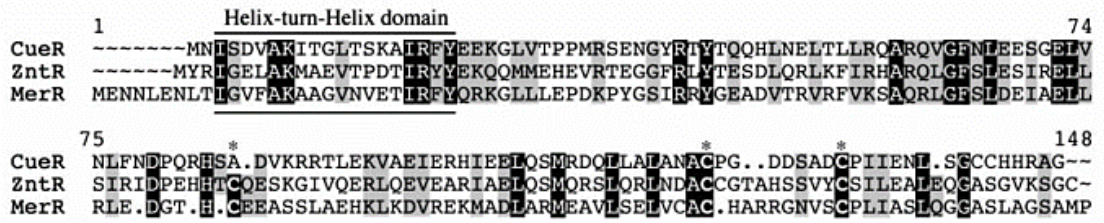

Figure 1.2-6: Amino acid alignment of ZntR and CueR from *E. coli* and MerR from Tn501. Identical residues are highlighted in black, similar residues in grey. The conserved cysteine residues are shown with an asterisk (reprinted modified with permission from Outten et al. 2000, copyright 2000 American Society for Biochemistry and Molecular Biology).

1.2.3.2.1 MerR

The MerR dimer regulates transcription of the *mer* genes under the $P_{merTPAD}$ promoter. In its apo form it forms a closed complex with the RNA polymerase and the promoter. This state leaves the -35 region accessible while the -10 site is blocked by MerR. Consequently the repression of the *mer* genes is somewhat leaky (Kulkarni & Summers 1999). MerR has been shown to bind Hg(II), inducing a conformational change of MerR and transcription of the *mer* genes (Helmann et al. 1990; Utschig et al. 1995). The induction is ultrasensitive, showing little increase (10%) from the background activity for Hg(II) concentrations of up to 3.6 µM Hg(II) and a rapid increase to 90% maximum activity at 32 µM Hg(II). Maximal induction with no further response to increased Hg(II) levels was observed from 100 µM Hg(II) (Ralston & O'Halloran 1990).

Recent results indicate that in addition to the aforementioned mechanism, the *in vitro* dissociation rate of MerR from DNA increases in response to Hg(II) at concentrations of 5 – 10^4 µg/L (approx.. 24 nM – 48 µM) (Taniguchi et al. 2014). It has previously been proposed that this increased dissociation rate might aid the

replacement of Hg(II) bound MerR-dimers (HG(II)-$MerR_2$) by non-metal-bound $MerR_2$, reducing expression faster after the Hg(II) stimulus is removed (Brown et al. 2003). MerD, one of the proteins of which expression is controlled by $P_{merTPAD}$, seems to support this replacement mechanism (Hobman et al. 2005). Apart from Hg(II), studies have shown *in vitro* binding of Cd(II) and Zn(II) at a sensitivity that is 10^2- and 10^3-fold lower than for Hg(II) with similarly ultrasensitive response curves (Ralston & O'Halloran 1990).

The promoter region of the *mer* operon (figure 1.2-6) has been the subject of several studies. It has been shown that small deletions in the 19 bp spacer between the -35 and -10 sequences lead to upregulation of the promoter and loss of normal activation through MerR (Lund & Brown 1989a; Lund & Brown 1989b; Parkhill & Brown 1990). Expanding the spacer by 1 or 2 bp removes promoter activity and changes to the imperfect inverted repeat binding site affect the regulation through MerR but not the constitutive promoter activity (Park et al. 1992; Livrelli et al. 1993).

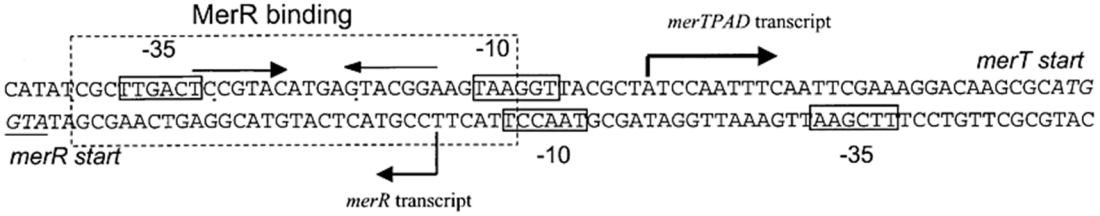

Figure 1.2-7: Structure of the promoter/operator region of the *mer* operon of Tn*501*.
The -35 and -10 regions of the promoters $P_{merTPAD}$ and P_{merR} are boxed and labelled against the coding strands of the appropriate promoters. The MerR binding sequence is boxed according to DNA footprinting and the symmetrical binding sequence is indicated by arrows (reprinted with permission from Brown et al. 2003, copyright 2003 Federation of European Microbiological Societies).

1.2.3.2.2 ZntR

ZntR was the first *E. coli* metal-dependent MerR-like regulator to be described and was originally named YhdM (Brocklehurst et al. 1999). The 141 amino acid ZntR protein is similar to MerR (Christie et al. 1994). In its dimeric form it exhibits an activator/repressor function based on DNA distortion similar to that of MerR and CueR (Outten et al. 1999; Philips et al. 2015). ZntR binds to P_{zntA}, which includes the characteristically elongated sequence between the -35 and -10 motifs seen in MerR-

family proteins. In contrast to $P_{merTPAD}$ and P_{copA}, this spacer is 20 bp long (Brocklehurst et al. 1999; Outten et al. 1999). This suggests that the deformation of P_{zntA} by ZntR must be greater than that of $P_{merTPAD}$ and P_{copA} by their respective trans elements, as all of these promoters are recognised by the σ^{70} subunit of the RNA polymerase (Hobman et al. 2005; Philips et al. 2015). The metal specificity of ZntR is not fully understood but is believed to be based on five cysteine residues in a principle similar to that seen in MerR (Outten et al. 1999; Khan et al. 2002; figure 1.2-4, page 31).

In vivo ZntR controlled expression is activated by Zn(II), Cd(II) and Pb(II). Timed induction appears to be strongest by Zn(II) and was shown over a range of 10^2-10^3 µM (Brocklehurst et al. 1999). Continuous induction throughout growth has indicated a stronger induction for 10^2 µM Cd(II) and also shows activation through Pb(II) (Binet & Poole 2000).

The activation of P_{zntA} by Zn(II)-$ZntR_2$ occurs with a Hill coefficient of 3.2, indicating a hypersensitive biological switch (Brocklehurst et al. 1999; Brown et al. 2003 ;figure 1.2-7 C). The gene controlled by ZntR is *zntA*, encoding an ATPase transporting Zn(II), Cd(II) and Pb(II) out of the cell (Brocklehurst et al. 1999; Outten et al. 1999). While no direct repressing function of ZntR has been shown so far, recent results for another MerR family member, CueR, make it likely that the mechanism in ZntR works in a similar way (see 1.2.3.2.3) (Hobman et al. 2005).

1.2.3.2.3 CueR

CueR has been shown to regulate the activity of the promoters P_{copA} and P_{cueO} in *E. coli* in response to cellular copper levels, hence the name standing for Cu efflux regulator (Outten et al. 2000; Stoyanov et al. 2001; Changela et al. 2003; figure 1.2-8).

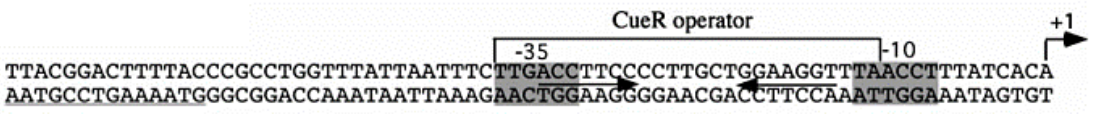

Figure 1.2-8: Promoter sequence of *copA*.
The palindrome is shown with arrows and the -35 and -10 regions of the promoter are highlighted in grey. The transcriptional start is marked by a cornered arrow and +1. The binding region of the CueR operator is marked by an accordingly labelled brace above the sequence .

Recently the allosteric transcriptional regulation mechanism has been resolved. The previously described mechanism of DNA distortion in the MerR family was shown with CueR in its activator and repressor state bound to the core region of P_{copA}. Like MerR and ZntR, CueR is active as a dimer and functions as an expression regulator through interactions with the promoter DNA and the RNA polymerase. In its repressor state it prevents contact between the polymerase and the -10 element of the promoter by forcing the DNA to bend. In its metal bound shape this bend is resolved and the DNA distorted to present the -35 and -10 regions to the polymerase in an ideal conformation (figure 1.2-9; Philips et al. 2015). This presents the first clear evidence that MerR is not the only protein in the MerR family that exhibits an active repression activity in addition to the widely accepted activator activity. While the work of Philips et al. indicates that MerR-like proteins (CueR) stay bound to the DNA with similar affinities regardless of metal presence, Taniguchi et al. have shown that the addition of Hg reduces the affinity of MerR for its binding site in $P_{merTPAD}$ as it had previously been proposed by Brown et al. (Brown et al. 2003; Taniguchi et al. 2014; Philips et al. 2015). It has yet to be confirmed whether either or both findings are influenced by experimental techniques or other factors and if the *in vivo* function of these proteins makes use of varying dissociation rates similar to the members of the SmtB-ArsR family.

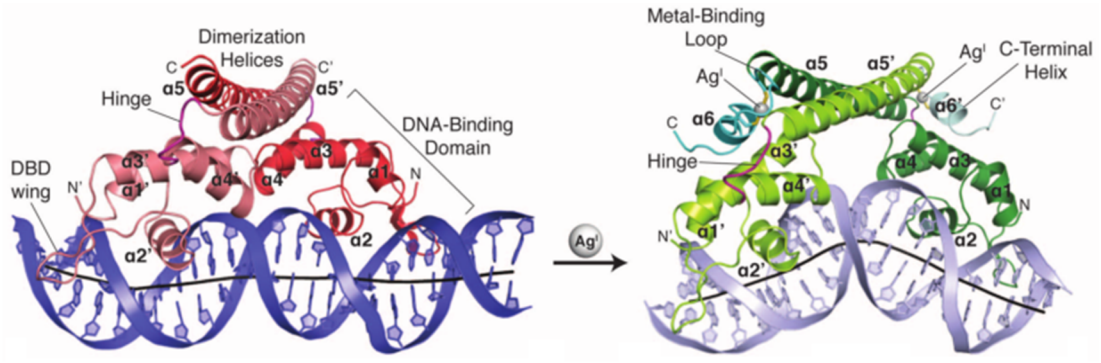

Figure 1.2-9: Crystal structures of the repressor and activator complex of CueR.
Shown are the core region of the promoter DNA with the activator (green) and repressor (red) states of CueR bound. The secondary structures, characteristic functional domains of CueR and the binding sites of Ag(I) to the protein are labelled accordingly (reprinted modified with permission from Philips et al. 2015, copyright 2015 American Association for the Advancement of Science).

CueR has been shown to react not only to Cu but also to Ag and Au (Outten et al. 2000; Stoyanov et al. 2001; Changela et al. 2003). In a series of *in vitro* experiments, Changela et al. showed the very high affinity with which Cu(I) binds to CueR *in vitro*, indicating a hypersensitive response with half-maximal transcription being observed at 0.7 ± 0.2 µM total Cu^+, at 0.013 ± 0.009 µM total Ag^+ and at 0.6 ± 0.3 µM total Au^+ (Changela et al. 2003). The high binding affinity is likely to be of importance for metal sensing in the cell, as metalloregulatory proteins effectively have to outcompete other intracellular metal ligands (Hobman et al. 2005; Giedroc & Arunkumar 2007).

Outten et al. showed that $CuSO_4$ levels above 1 µM in defined media lead to an increased *in vivo* expression of plasmid borne *lacZ* under the control of P_{copA}. Highest expression levels of *lacZ* in this setup were achieved with copper concentrations around 500 µM. No such induction was evident under equivalent zinc or mercury concentrations (Outten et al. 2000). In similar experiments, Stoyanov et al. showed induction by $CuSO_4$ and $AgNO_3$. In contrast to Outten and colleagues, Stoyanov et al. found the copper level to be required to achieve maximum induction to be much lower, around 50 µM (Stoyanov et al. 2001). This is likely to be due to differences in the experimental setup. Outten et al. used a single copy vector system carrying P_{copA} (Simons et al. 1987; Outten et al. 2000) while Stoyanov and

colleagues used a low-copy-number plasmid to carry the promoter sequence and a medium copy number plasmid to supply the regulating protein CueR (Stoyanov et al. 2001). However, total expression levels of the *lacZ* gene used in both studies seem to differ greatly. Stoyanov's evidently sensitive screening method did not show any *in vivo* induction by Ni, Cd, Hg, Pb or Zn (Stoyanov et al. 2001).

1.2.4 Biosensors for heavy metals

On the following pages a brief collection of exemplary biosensors for heavy metals that follow multiple strategies will be presented. Where possible these examples have been chosen by actuality at the time of writing, relevance for this work and to cover a wide range of approaches.

1.2.4.1 Whole-cell sensors based on genetically engineered microorganisms

1.2.4.1.1 A pH sensor for As

An example of a biosensor currently in the advanced stages of development for field use, is the *Bacillus subtilis* spore based sensor for As in drinking water. In its core, this sensor relies on the principles of the Edinburgh Arsenic Biosensor (Aleksic et al. 2007; Joshi et al. 2009; de Mora et al. 2011). The latter is based on a pH change in the surrounding medium due to arsenite induced (P_{ars}) expression of β-galactosidase in *E. coli* and the resulting lactose fermentation. Over the years, improvements in media formulation and screening methods have allowed this sensor to become sensitive enough to detect arsenate levels below 10 ppb (10 µg/L or approx. 0.13 µM) in an over-night incubation, in principle fulfilling the sensing needs in regard to drinking water quality (de Mora et al. 2011). In the past 5 years, a similar, RFP based system has been adopted for use in *B. subtilis*, with the aim of overcoming the inherent problems of shelf-life and legislative regulation of potential GEM release. Briefly, the *B. subtilis* version of the sensor incorporates a chromosomal copy of the sensing mechanism in a knock-out mutant unable to survive in the wild for any prolonged period of time. Being a spore forming bacterium, *B. subtilis* has the advantage of dramatically increasing the shelf life of the sensor from a few weeks to years (de Mora et al. 2011; David Radford, private communication). However, the legislative restrictions for the field use of such a sensor remain a limiting factor and as of today it is uncertain whether permission for release will be granted.

1.2.4.1.2 A luminescence sensor for Cu, Ag and Au

In 2003, Stoyanov et al. measured the cytoplasmic levels of Cu, Ag and Au in *E. coli* with a biosensor based on P_{copA} and the metal-responsive regulator CueR (see 1.2.3,

for an introduction to metal-responsive regulators and 1.2.3.2.3 for details on P_{copA}/CueR). To show the intracellular levels of the aforementioned metals, they placed the *Vibrio fischeri lux* gene cluster under the control of P_{copA} in a high copy number plasmid. Their initial results showed that a knock out mutant of *E. coli*, lacking a functional Cu and Ag efflux pump CopA was more sensitive to metal levels in the surrounding medium. With this sensor system they were able to detect Ag in the range between 0.01 µM and 2 µM, Cu between 0.5 µM and 250 µM and Au between 10 µM and 60 µM (Stoyanov et al. 2003). These detection levels are suited to sense Cu in the relevant range around 15 µM to 30µM and Ag around 0.9 µM. The practicality as an actual sensor might be limited, due to the susceptibility of luciferase based systems to drastic output changes as a consequence of variations in experimental timing and ambient temperature.

1.2.4.2 Protein based cell-free biosensors

1.2.4.2.1 *A cis-trans* sensor for As and Cd

Kawakami et al. described a protein based cell-free sensor based on *cis-trans* interactions, that is the interaction between a regulatory DNA binding protein body, the trans element, and its binding site on the DNA, the cis element (Kawakami et al. 2010). For this sensor ArsR from *E. coli* and CadC, a Cd and Pb reactive *trans* element from *Staphylococcus aureus,* were each fused to a green fluorescent protein (GFP) and expressed in *E. coli*. Cell extracts containing the fusion protein were incubated in microplate wells that contained multiple well-bound copies of the promoter/operator regions known to be controlled by the respective native proteins *in vivo*. Consistent with the known DNA binding and unbinding mechanism of SmtB-ArsR family proteins, it was found that fusion-proteins from cell extracts that had been pre-incubated with corresponding heavy metals would not bind the immobilised DNA and be readily removed by a wash step. Such extracts that were free from heavy metals would allow the contained fusion-proteins to attach to their respective binding sites on the DNA molecules and withstand being washed away. This effect was found to be proportional to the heavy metal concentrations used. As a consequence the fluorescence intensity of the remaining fusion proteins in each well

after washing was inversely proportional to the total heavy metal concentration in sample waters used in the pre-incubation (Kawakami et al. 2010). In a later modification of this sensor the fusion proteins were pre-incubated with the immobilised DNA in a reaction vessel and a sample mixture added. After incubation of 15 – 30 minutes the supernatant from the reaction vessel was removed and measured for fluorescence. In this case the measured fluorescence was found to be directly proportional to the heavy metal concentration in the sample mixture as a consequence of the release of *trans* factors from their respective *cis* elements.
The successfully measured heavy metal concentrations were as low as 1 µg/L for Cd and 5 µg/L for As(III) in purified water and around 10 µg/L for both metals in tap water and bottled mineral water. Similar sensitivities were found when the *cis-trans* complexes were freeze dried and stored at 4°C (Siddiki et al. 2011).

In their testing of various waters, Siddiki et al. experienced an influence of water hardness, namely $CaCO_3$ content, on the sensor reaction, shifting the response pattern (Siddiki et al. 2011). The influence of factors like this on sensor performance under field conditions can often turn an easy to use device into a laborious impracticality through the necessity for calibration standards, pre-testing or pre-treatment steps (Siddiki et al. 2011). Despite these problems and with some further experimental fine tuning, Siddiki et al. were able to measure As, Cd, Pb and Zn concentrations in diluted milk and yoghurt, demonstrating the overall robustness of the system (Siddiki et al. 2012).

1.2.4.2.2 A *cis-trans* Hg sensor based on surface plasmon resonance

As aforementioned, Taniguchi et al. have recently shown that the dissociation rate of MerR is influenced by presence of Hg (Taniguchi et al. 2014). Their first approach was in its core identical to the previously described *cis-trans* sensor for As and Cd. However, changes in fluorescence activity for a system based on MerR-GFP – $P_{merTPAD}$ affinity changes in response to Hg were minimal compared to the previously described results. Suspecting the sensitivity of the fluorescence measurement to be insufficient for monitoring of the dissociation events, they employed surface plasmon resonance instead. With this significantly more sensitive method they were able to

show a change in dissociation rate that was linearly proportional to log[HgII] concentrations between 5 and 10^4 µg/L (approx.. 24 nM – 48 µM). While the use of surface plasmon resonance implies a significant increase in instrument cost over fluorescent measurement, this strategy bears several advantages too. Firstly the detection limit is astonishingly low and secondly they were able to shorten the essay time to 6 minutes, all while employing a reusable sensor chip with immobilised double stranded DNA including the corresponding *cis*-element. As in their previous experiments, Taniguchi et al. used mineral water to assess the influence of commonly found drinking water contents on the essay and found no detrimental effects on the system (Taniguchi et al. 2014). This sensor system shows how the use of highly sensitive instruments can greatly improve the overall sensitivity of a biosensor for laboratory use at the cost of field applicability and affordability.

1.2.4.2.3 DNA distortion detection based sensor for Hg and other heavy metals

Since it is not fully understood whether the proteins of the MerR family exhibit a DNA binding-unbinding activity in response to heavy metal presence, expression free sensors based on these proteins have been designed following a different strategy.

Chen et al. and Wegner et al. have described sensors that utilise the unique distortion mechanism of MerR-like proteins (Chen & He 2003; Wegner et al. 2007). In their approach they incorporated pyrene dyes into the separate strands of the duplex DNA of the promoter-operator region of MerR. Pyrenes form excited-state dimers, so called excimers. Since the distance and geometry between two pyrene units has a great influence on the Stokes shift, DNA distortion can be measured by observing this shift. Wegner et al. observed an emission maximum after excitation at 350 nm at around 480 nm while MerR was bound to the DNA in its metal-free state. Addition of Hg led to a dose dependent decrease in emission at 480 nm and increased emission at 380 nm (figure 1.2-14).

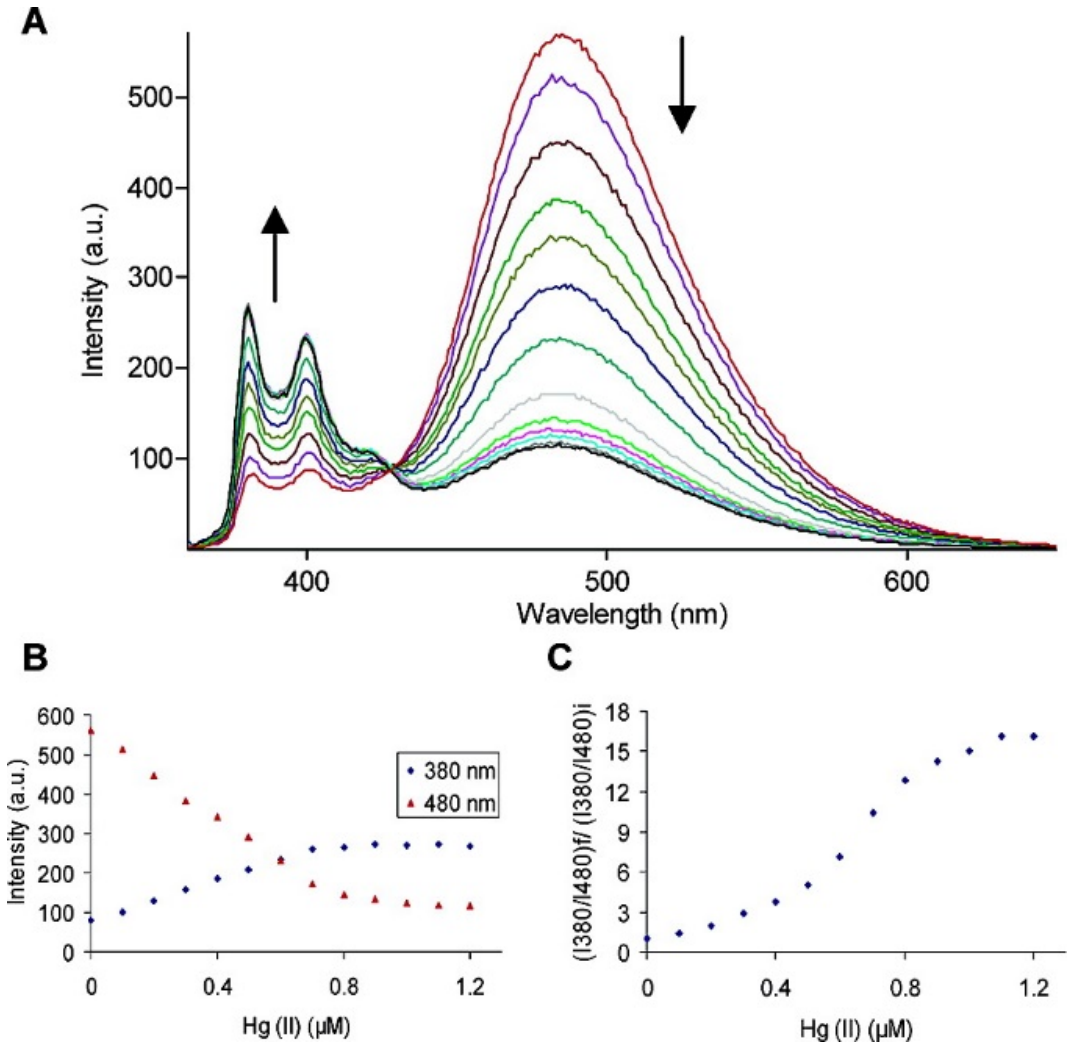

Figure 1.2-10: Fluorescence shift in response to DNA distortion.

A) fluorescence response of $MerR^2$-DNA complex (1 µM) to the addition of Hg^{2+} at 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, and 1.2 µM. The measurements were performed at room temperature in a nitrogen-purged buffer containing 10 mM Tris-HCl (pH 7.4), 100 mM potassium glutamate, 2 mM $MgCl_2$ and 5% glycerol. B) fluorescence intensity as a function of Hg^{2+} concentration. C, ratiometric calibration curve final I_{380}/I_{480} ratio over the initial I_{380}/I_{480} as a function of Hg^{2+} concentration (reprinted with permission from Wegner et al. 2007, copyright 2007 American Chemical Society).

With a detection limit of around 0.1 µM and a reasonable dynamic range up to around 1.1 µM, the sensor is relatively sensitive when compared to whole cell systems. Impressively the fluorescence shift at 1 µM Hg was strong enough to be observed by eye under UV light. The reaction was found to be highly specific for Hg

and very little activity was seen for any other metal added at similar or higher concentrations (figure 1.2-15; Wegner et al. 2007).

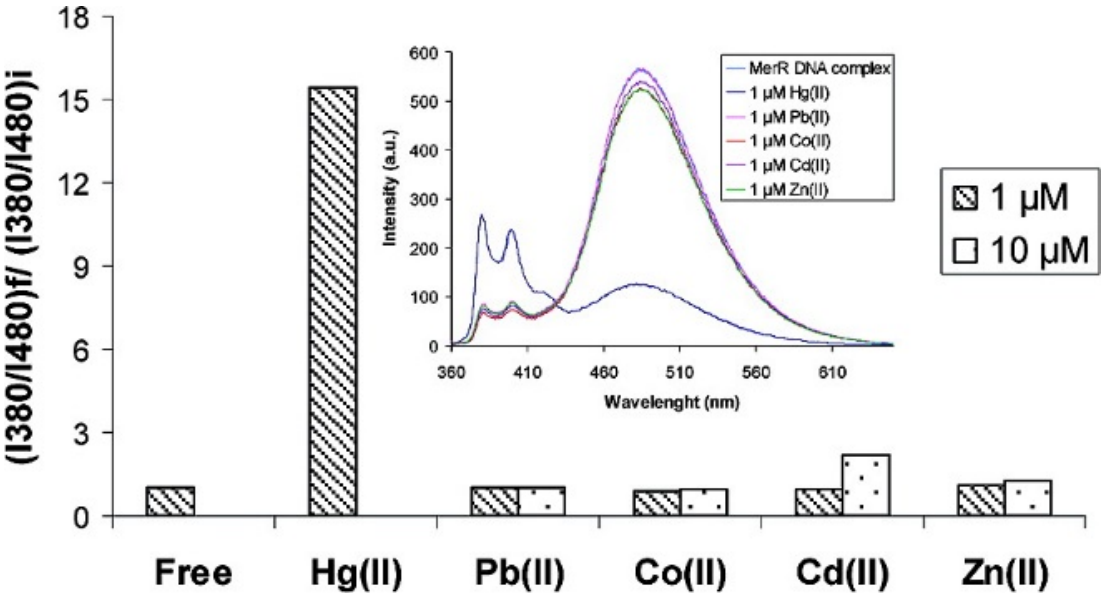

Figure 1.2-11: Ratiometric response for different metal ions at 1 and 10 µM.
The inset shows the fluorescence response of $MerR^2$-DNA complex to the addition of 1 µM of different metal ions. Measurement conditions are as above (reprinted with permission from Wegner et al. 2007, copyright 2007 American Chemical Society).

An advantage of this sensor over many other systems described is its ratiometric response. By measuring the ratio between two fluorescence signals from the same source and deriving the heavy metal concentration from this ratio, influences on the overall fluorescence brightness such as temperature and ionic strength have less impact on the measurement accuracy. While this gives the sensor a degree of robustness, the measurement in relatively specific nitrogen-purged buffer and the current lack of data for the influence of pollutants likely to be found in a real-life drinking water sample on the essay, leave some questions for the field-applicability of this otherwise very elegant approach (Wegner et al. 2007; Siddiki et al. 2011; Siddiki et al. 2012).

1.2.4.2.4 A multi metal sensor based on cell-free expression systems

In 2015, after work had started on the cell-free heavy metal sensing array presented in this work (3), the Bielefeld iGEM team presented a sensor based on similar principles for the detection of heavy metals and date rape drugs in water. Their aim was to develop a paper-based cell-free expression based sensor using various microbial heavy metal inducible promoters. Initially sensors were designed for As, chromium (Cr), Cu, Pb, Hg and nickel (Ni). The basic principle of all these sensors is the same with the exception of As. Superfolder GFP (sfGFP) was brought under the control of a metal inducible promoter on one plasmid. On a second plasmid, the regulatory protein was expressed under the control of a strong constitutive promoter. In the case of As, the second plasmid was omitted and the arsenic responsive operator region of P_{ars} was placed behind a strong T7 promoter. It was hoped that the approach of including sensing and regulating components on two separate plasmids would allow for adjustment of the ratio at which the plasmids are provided and expression was likely to take place, thus ideally allowing an adjustment of the sensors' characteristics to the sensing needs. Out of the 7 constructs for the metals described above only two – Cu and Hg – showed satisfactory performance in the cell-free expression systems. These two sensors were later tested on paper as a carrier material and were found to be sufficiently active to be used for assessment of Cu and Hg in water while their activity was monitored with the help of a smartphone camera. However, detection limits remain somewhat unclear.

1.3 Aims and objectives

Heavy metals can be a major threat to human health. In the past, industrial pollution or ill-fated attempts at developmental aid have exposed hundreds of millions of people to heavy metals in drinking water. In the case of arsenic, this problem is ongoing and no solution seems to be at hand. Relatively low cost filters for purification of As polluted waters are available but require maintenance and control of their effectiveness. As such, frequent assessment of the water quality of each tube well is currently the only practical measure against this public health crisis. Since the number of tube wells is vast, a simple test that can be performed by each tube well user, free from legislative restrictions, is required (Smith et al. 2000; de Mora et al. 2011). The development of such a sensor approach was one aim of this work. See chapter 3 for more details on this sensor and its aims.

Apart from exclusive arsenic sensing, a multiplexed sensor adding detection abilities for cadmium, copper, lead, mercury, silver and zinc (and as a by-product gold) might be advantageous for water testing in a multitude of situations, if detection times are reasonably short and detection limits are sufficiently low (Chen & Rosen 2014). As many metalloregulatory proteins detect several metals, multiple sensors working together can elucidate which specific metal is being detected. As such, the further aims of this project were to develop a prototype set of cell-free biosensors, allowing for the specific detection of heavy metals through a mutually exclusive response. See chapters 4 through 6 for details on these sensor systems and the associated work and approaches.

The key aspects of the work presented in chapter 3 were:

1. To design and produce a translational ArsR fusion protein and assess its interaction with the P_{ars} promoter.
2. To demonstrate the As(III) responsiveness of this interaction.
3. To visualise the As(III) dependent dissociation of the fusion protein from a P_{ars} derived probe on a solid support.

The key aspects of the work presented in chapter 4 were:

1. To design and construct proof of concept biosensors for As, Ag, Au, Cd, Cu, Hg, Pb and Zn based on the previously introduced metal responsive promoters P_{ars}, P_{copA}, $P_{merTPAD}$, P_{zntA} and their regulatory protein partners.
2. To assess their activity *in vitro* in a cell free expression system.
3. To investigate the potential of a tuneable As sensor.

The key aspects of the work presented in chapter 5 were:

1. To establish the production of in house cell extracts for cell free protein expression.
2. To improve extract performance to allow larger scale testing of the biosensors created.

The key aspects of the work presented in chapter 6 were:

1. To incorporate the fluorescent RNA aptamer Spinach2 into the biosensor concepts presented in chapter 4.
2. To test *in vivo* and i*n vitro* characteristics of Spinach2 and the sensors combined with it.
3. To identify the minimal requirements for Spinach2 transcription *in vitro* as a first step towards a minimal, high speed cell free sensor approach.

2 Materials and methods

2.1 Materials

2.1.1 List of Laboratory suppliers

Unless otherwise stated, all materials used were acquired from the following suppliers, their abbreviated names given below will be used for reference throughout this work.

Table 2.1-1: List of laboratory consumable suppliers

Abbreviated name	Full supplier name	Detail
EMD Millipore	EMD Millipore (U.K.)	Feltham, UK
Invitrogen	Thermo Fisher Scientific	Paisley, UK
Merck	Merck Serono	Feltham, UK
Merck	Merck Chemicals	Beeston, UK
NEB	New England Biolabs	Ipswich, MA, USA
Promega	Promega UK	Southampton, UK
Sigma	Sigma-Aldrich	Gillingham, UK
Thermo Fisher	Thermo Fisher Scientific	Paisley, UK
Bio-Rad	Bio-Rad Laboratories	Hemel Hempstead, UK
GE Healthcare	GE Healthcare	Little Chalfont Bucks, UK
Expedeon	Expedeon	Swavesey, UK
Brand	BrandTech Scientific	Essex, CT, USA
Greiner	Greiner bio-one	Stonehouse, UK
BMG	BMG Labtech	Ortenburg, Germany
Tocris Bioscience	Tocris Bioscience	Bristol, UK

2.1.2 Bacterial strains, Plasmids and Oligonucleotides

Table 2.1-2: List of *E.coli* strains.

Strain	Details	Use	Source
BL21 (DE3)	F⁻ *ompT gal dcm lon hsdS$_B$*($r_B^- m_B^-$) λ(DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*]) [*malB*$^+$]$_{K\text{-}12}$(λ^S)	Recombinant protein expression.	Lab Stock
DH10β	F⁻ *endA1 deoR*$^+$ *recA1 galE15 galK16 nupG rpsL Δ(lac)X74* φ80*lacZΔM15 araD139 Δ(ara,leu)7697 mcrA Δ(mrr-hsdRMS-mcrBC)* Str^R λ(DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*])	General cloning, except constructs containing streptomycin or ampicillin resistance genes.	Lab Stock generated from NEB commercial competent cells.
JM109 (DE3)	endA1 glnV44 thi-1 relA1 gyrA96 recA1 mcrB$^+$ Δ(lac-proAB) e14- [F' traD36 proAB$^+$ lacI^q lacZΔM15] hsdR17($r_K^- m_K^+$) λ(DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*])	General cloning for constructs containing streptomycin or ampicillin resistance.	Lab Stock
MG1655	K-12 F⁻ λ⁻ *ilvG⁻ rfb-50 rph-1*	Genomic DNA preparations and as PCR template.	Lab Stock
Top10	F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(Str^R) endA1 λ$^-$	General cloning, except constructs containing streptomycin or ampicillin resistance genes.	Lab Stock generated from NEB commercial competent cells.

Table 2.1-3: List of plasmids.

Plasmid	Length	Relevant features	Source
pSB1A2	2079 bp	High copy number, ampicillin resistance	Registry for standard biological parts
pSB1C3	2070 bp	High copy number, chloramphenicol resistance	Registry for standard biological parts
pET28 a (+)	5369 bp	High copy number, kanamycin resistance, C- and N-terminal His tags, T7 promoter, *lac* operator	EMD Millipore
pBEST-*luc*	4486 bp	High copy number, ampicillin resistance, tac promoter, *lac* operator, firefly luciferase gene	Promega
pGEM-T Easy	3000 bp	Direct cloning of PCR constructs with poly A overhangs, T7 promoter, blue/white selection capable	Promega

Table 2.1-4: List of oligonucleotides.

Name	Sequence 5' – 3'	Notes
ArsR-BS-for	ggttacttacacattcgttaagtcatatatgtttttgacttatccgcttcg	ArsR binding site
ArsR-BS-polyTC-fw	ttttttttttccccccccccttacacattcgttaagtcatatatgtttttgacttatcc	
ArsR-BS-rev	aacccgaagcggataagtcaaaaacatatatgacttaacgaatgtgtaagt	
ArsR-BS-TC-bw	ggataagtcaaaaacatatatgacttaacgaatgtgtaag	
arsR-G3 bw	tccgcccgagccaccgccaccagaactgcaaatgttcttactgtccc	MABEL
DF_G3-linker	gaggtgggtcgggtggcggcggatca	PaperClip downstream primers for clip assembly and amplification.
DF_*EcArsR*	acaaaactgttccggggacagtaagaacatttgcagttaa	
DF_*EcArsR*_ns	tcgacaaaactgttccggggacagtaagaacatttgcagt	
DF_*EcCadC*_ns	ttatgttgtaccttatctcgacaaatttcttgcttcagaa	
DF_*EcCueR*_ns	tatcgaaaatctctccggctgctgtcatcatcgggcaggg	
DF_*EcMerR*_ns	actacagggcgaagcaggcctggcaaggtcagctatgcct	
DF_*EcZntR*_ns	agctcttgaacaaggggcgagtggcgttaagagtggttgt	
DF_*lacZ*	ggtcgctaccattaccagttggtctggtgtcaaaaataa	
DF_*luc*	cctcataaaggccaagaagggcggaaagtccaaattgtaa	
DF_*luc*_ns	gatcctcataaaggccaagaagggcggaaagtccaaattg	
DF_*mCherry*	ccactccaccggcggcatggacgagctgtacaagtaataa	
DF_*mCherry*_ns	gggccgccactccaccggcggcatggacgagctgtacaag	
DF_P_{ars}_sRBS	cgcttcgaagagagacactacctgcaatacggttactctc	
DF_pBEST	atcagccagaaaaccgaattttgctgggtgggctaacgat	
DF_pBEST-P_{lac}	aattgtgagcggataacaatttcacacaggaaacaggatc	
DF_P_{copA}_sRBS	tttatcacagccagtcaaaactgggtaacctacggttaaa	

Name	Sequence 5' – 3'	Notes
DF_pET28 C-His	tcccctctagaaataattttgtttaactttaagaaggaga	
DF_pET28_sRBS	tctagaaataattttgtttaactttactacggttattcac	
DF_$P_{merTPAD}$_sRBS	ctatccaatttcaattcgaatctacgcctca	
DF_$P_{merTPAD}$-WT	agtaaggttacgctatccaatttcaattcgaaaggacaag	
DF_P_{zntA}_sRBS	tgtatccttcggttaatgagaaaaaaccactacggaactc	
DF_Spinach2	cggccgcgggtccagggttcaagtccctgttcgggcgcca	
DR_(GGGGS)3-linker	ggctgatccgccgccacccgaccc	
DR_*EcArsR*	ggcttaactgcaaatgttcttactgtccccggaacagttt	
DR_*EcArsR*_ns	ggcactgcaaatgttcttactgtccccggaacagttttgt	
DR_*EcCueR*_ns	ggcccctgcccgatgatgacagcagccggagagattttcg	
DR_*EcMerR*_ns	gccaggcatagctgaccttgccaggcctgcttcgccctgt	
DR_*EcZntR*_ns	ggcacaaccactcttaacgccactcgccccttgttcaaga	
DR_*lacZ*	ggcttatttttgacaccagaccaactggtaatggtagcg	
DR_*luc*	ggcttacaatttggactttccgcccttcttggcctttatg	
DR_*luc*_ns	ggccaatttggactttccgcccttcttggcctttatgagg	
DR_*mCherry*	ggcttattacttgtacagctcgtccatgccgccggtggag	
DR_*mCherry*_ns	ggccttgtacagctcgtccatgccgccggtggagtggcgg	
DR_P_{ars}_sRBS	ggcgagagtaaccgtattgcaggtagtgtctctcttcgaa	
DR_pBEST	ggcatcgttagcccacccagcaaaattcggttttctggct	
DR_pBEST-P_{lac}	ggcgatcctgtttcctgtgtgaaattgttatccgctcaca	
DR_P_{copA}_sRBS	ggctttaaccgtaggttacccagttttgactggctgtgat	
DR_pET28 C-His	ggctctccttcttaaagttaaacaaaattatttctagagg	
DR_pET28_sRBS	ggcgtgaataaccgtagtaaagttaaacaaaattatttct	
DR_$P_{merTPAD}$_sRBS	ggctgaggcgtagattcgaattgaaattggatagcgtaac	
DR_$P_{merTPAD}$-WT	gcgcttgtcctttcgaattgaaattggatagcgtaacctt	
DR_P_{zntA}	ggcgagttccgtagtggtttttctcattaaccgaaggat	
DR_Spinach2	ggctggcgcccgaacagggacttgaaccctggacccgcgg	
EcArsR f BspHI	gtttcatgatgtcatttctgttacccatc	Restriction cloning.
EcArsR f NdeI	gttcatatgtcatttctgttacccatcc	
EcArsR r no stop	actgcaaatgttcttactgtccc	
EcArsR-m*Cherry/EYFP* f	gggacagtaagaacatttgcagtatggtgagcaagggcg	
EcArsR-mCherry/EYFP r	cgcccttgctcaccatactgcaaatgttcttactgtccc	
Ec P_{ars} f	ccaactcaaaattcacacctattacc	Amplification primers.
E.coli P_{ars} r	attgcgctcctgattgttg	
*EcArsR*_bw_ns	actgcaaatgttcttactgtccc	
EYFP/ mCherry f	atggtgagcaagggcg	
EYFP/mCherry r	ttattacttgtacagctcgtccatg	
EYFP/mCherry r ns	cttgtacagctcgtccatgc	
G3-*mCherry* fw	ggtgggtcgggtggcggcggatcaatggtgagcaagggcg	Linker
His_fw	caccaccaccaccacc	Sequencing
*luc*_fw	atggaagacgccaaaaac	Amplification
M13_fw_17	gttttcccagtcacgac	Sequencing
M13_fw_24	cgccagggttttcccagtcacgac	
M13_rev_17	caggaaacagctatgac	
M13_rev_24	tcacacaggaaacagctatgac	
mCherry/EYFP BB f	gtagaattcgcggccgcttctagatggtgagcaagggcg	Restriction cloning.
mCherry/EYFP BB r	gaactgcagcggccgctactagtattattacttgtacagctcgtccatg	
mCherry/EYFP ns XhoI r	gatctcgagcttgtacagctcgtcc	

Name	Sequence 5' – 3'	Notes
mCherry/EYFP XhoI r	gttctcgagttattacttgtacagctcgtc	
*mCherry*_fw	atggtgagcaagggc	Amplification
P_{copA}_bw	aaaacactcctttaagacagttttgactggctgtg	
P_{copA}_fw	gccgatgcgggaggtaattcc	
pET28_sRBS_bw	gtgaataaccgtagtaaagttaaacaaaattatttctagagg	MABEL
pET28-*arsR*_sRBS_fw	gccatgtcatttctgttacc	
$P_{merTPAD}$_bw	gcgcttgtcctttcgaattgaaattggatagcg	Amplification
P_{zntA}_bw	ggcatcctccggttaagttttttctcattaaccgaagg	
P_{zntA}-fw	gattgctgcggcctg	
*sce*_fw	atgccgtctctgagatttc	
sce-seqE2-bw	gatcgccgagcttcc	Sequencing
sce-seqE2-fw	gcgtgaagaagtggatcac	
sce-seqE-bw	cccccagacgaaattgt	
sce-seqE-fw	accgcgacgtcct	
T7_promoter_f	taatacgactcactataggg	
T7_terminator_r	gctagttattgctcagcgg	
UF_(GGGGS)3-linker	gcctctggtggcggtggctcgggcg	Linker
UF_*EcArsR*	gccatgtcatttctgttacccatccaattgttcaaaattcttg	Upstream PaperClip primers for assembly of clips and part amplification.
UF_*EcCueR*	gccatgaacatcagcgatgtagcaaaaattaccggcctgacca	
UF_*EcMerR*	gccatggaaaataatttggaaaacctgaccattggcgtttttg	
UF_*EcZntR*	gccatgtatcgcattggtgagctggcaaaaatggcggaagtaa	
UF_*lacZ*	gccatgaccatgattacggattcactggccgtcgttttacaac	
UF_*luc*	gccatggaagacgccaaaaacataaagaaaggcccggcgccat	
UF_*mCherry*	gccatggtgagcaagggcgaggaggataacatggccatcatca	
UF_P_{ars}	gccccaactcaaaattcacacctattaccttcctctgcactta	
UF_pBEST	gccaatgtaactgtattcagcgatgacgaaattcttagctatt	
UF_P_{copA}	gccgatgcgggaggtaattcctcaccccggtgccgattttcag	
UF_pET28 C-His	gcccaccaccaccaccaccactgagatccggctgctaacaaag	
UF_$P_{merTPAD}$	gccatcgcttgactccgtacatgagtacggaagtaaggttacg	
UF_P_{zntA}	gccgattgctgcggcctgctactttgccggtcacttcctgatc	
UF_Spinach2	gccgcccggatagctcagtcggtagagcagcggccggatgtaa	
UF_Spinach2_forMER	cgcgcccggatagctcagtcggtagagcagcggccggatgtaa	
UR_G3-linker	acctccgcccgagccaccgccaccaga	
UR_*EcArsR*	gaattttgaacaattggatgggtaacagaaatgacat	
UR_*EcCueR*	tcaggccggtaatttttgctacatcgctgatgttcat	
UR_*EcMerR*	aaacgccaatggtcaggttttccaaattattttccat	
UR_*EcZntR*	cttccgccatttttgccagctcaccaatgcgatacat	
UR_*lacZ*	gtaaaacgacggccagtgaatccgtaatcatggtcat	
UR_*luc*	gcgccgggcctttctttatgtttttggcgtcttccat	
UR_*mCherry*	tgatggccatgttatcctcctcgcccttgctcaccat	
UR_P_{ars}	gtgcagaggaaggtaataggtgtgaattttgagttgg	
UR_pBEST	agctaagaatttcgtcatcgctgaatacagttacatt	
UR_P_{copA}	aaaatcggcaccggggtgaggaattacctcccgcatc	
UR_pET28c-His	tgttagcagccggatctcagtggtggtggtggtggtg	
UR_$P_{merTPAD}$	cttacttccgtactcatgtacggagtcaagcgat	
UR_P_{zntA}	caggaagtgaccggcaaagtagcaggccgcagcaatc	
UR_Spinach2	catccggccgctgctctaccgactgagctatccgggc	

2.1.3 DNA sequencing

Sanger DNA sequencing (Sanger et al. 1977) was performed by Edinburgh Genomics, Edinburgh, UK.

2.1.4 Molecular weight markers and DNA ladders

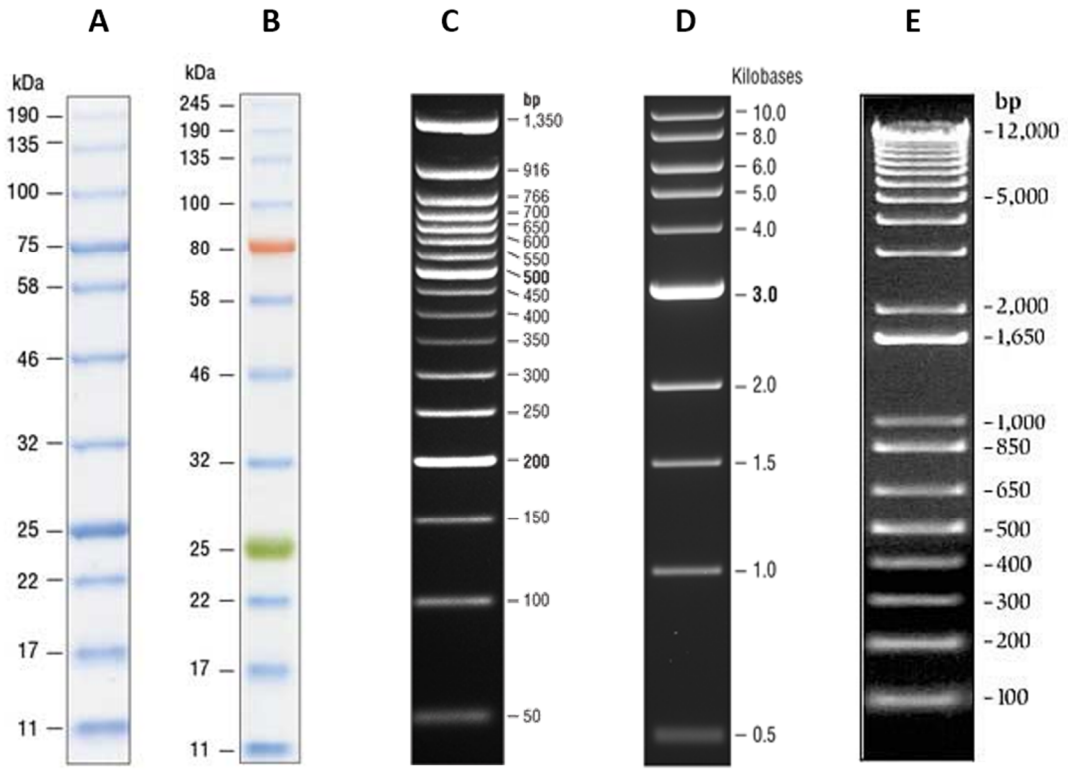

Figure 2.1-1: Protein molecular weight markers and DNA ladders used in this work.
A) Blue prestained protein standard broad range (11-190 kDa), 10-20% tris-glycine SDS-PAGE gel (New England Biolabs). B) Colour prestained protein standard, broad range (11-245 kDa) 10-20% tris-glycine SDS-PAGE gel (New England Biolabs). C) 50 bp DNA ladder, ethidium bromide staining on a 3% TBE agarose gel (New England Biolabs). D) Quick-Load 1 kb DNA ladder, ethidium bromide staining on a 0.8% TAE agarose gel (New England Biolabs). E) 1 kb Plus DNA ladder, ethidium bromide staining on a 0.9% TAE agarose gel (Thermo Fisher).

2.1.5 Enzymes

All restriction enzymes were supplied by NEB and used according to manufacturer specifications.

Table 2.1-5: DNA Polymerases.

Name	Notes	Source
GoTaq® Flexi	Direct to gel amplification buffer	Promega
GoTaq® G2 Flexi	Second generation GoTaq®	Promega
KOD Hot Start	Proof reading, superior performance in PaperClip assemblies and in high GC amplifications	EMD Millipore
NEB Taq	Cost effective	NEB
Phusion High-Fidelity	Proof reading	Thermo Fisher
Q5® High-Fidelity	Proof reading	NEB

2.1.6 Growth media

All growth media was autoclaved at 212°C for 15 min prior to use unless otherwise stated.

Table 2.1-6: Lysogeny broth after Bertani 1951.

Ingredient	Amount
Tryptone	10 g
Yeast Extract	5 g
NaCl	10 g
Agar (if required)	15 g
ddH_2O	to 1 L

Table 2.1-7: Terrific Broth after Tartoff & Hobbs 1987.

Ingredient	**Amount**
Tryptone	12 g
Yeast Extract	24 g
Glycerol 10% (v/v)	4 mL
ddH_2O	to 900 mL
Autoclave, then add the following as a sterile solution:	
KH_2PO_4	2.31 g
K_2HPO_4	12.54 g
ddH_2O	to 100 mL

Table 2.1-8: M9 medium after Sambrook et al. 1989.

Ingredient	**Amount / Adjustment**
Na_2HPO_4	6 g
KH_2PO_4	3 g
NaCl	0.5 g
NH_4Cl	1 g
Glucose 20% (w/v)	20 mL
ddH_2O	to 1 L

Table 2.1-9: Autoinduction medium.

Ingredient	**Amount / Adjustment**
Na_2HPO_4	6 g
KH_2PO_4	3 g
Adjust pH to 7.2 (HCl/NaOH)	
Tryptone	20 g
Yeast extract	5 g
NaCl	5 g
ddH_2O	to 947.5 mL
Autoclave, then add sterile:	
Glycerol 20% (v/v)	30 mL
Glucose 20% (w/v)	2.5 mL
Lactose 20% (w/v)	10 mL

Table 2.1-10: 2 x YPTG medium.

Ingredient	**Concentration**
Peptone	1.6% (w/v)
Yeast extract	1% (w/v)
NaCl	0.5% (w/v)
NaH_2PO_4	22 mM
Na_2HPO_4	40 mM
Glucose	100 mM

Table 2.1-11: List of Antibiotics.

Antibiotic	Solvent	Stock concentration	Working concentration
Ampicillin	ddH_2O	50 mg/mL	50 µg/mL
Carbenicillin	ddH_2O	100 mg/mL	100 µg/mL
Chloramphenicol	Ethanol	40 mg/mL	40 µg/mL
Kanamycin	ddH_2O	50 mg/mL	50 µg/mL*
* Kanamycin working concentrations were doubled to 100 µg/mL for M9 and autoinduction medium due to high phosphate concentrations inducing kanamycin resistance.			

2.1.7 Buffers

Buffers are included with their respective (main) applications.

2.2 Methods

2.2.1 DNA manipulation

2.2.1.1 Polymerase chain reaction

PCR amplification of DNA was performed using a range of polymerases as specified in section 2.1.5, page 52, table 2.1-5. Programs and reaction compositions were altered to suit individual requirements in each reaction, general procedures are listed below. Glycerol was added to amplifications of high GC templates. Annealing temperatures were adjusted in accordance with the lowest melting temperature (T_m) as calculated by the SnapGene software. PCR products were purified either by column purification (2.2.2.3, page 68) or gel purification (2.2.2.1 & 2.2.2.2, page 68). If minimal template carry over was desired, completed PCR reactions were supplemented with 1 µL of the restriction enzyme DpnI and incubated at 37°C for 1 h before gel purification.

Table 2.2-1: Standard PCR reactions.

Reaction component	**GoTaq Flexi, GoTaq G2 Flexi, NEB Taq**	**Phusion HF**	**Q5 HF**	**KOD hot start**
DNA template	20 – 200 ng	20 – 200 ng	20 – 200 ng	20 – 200 ng
Primer (10 mM)	0.5 µL each	1	2.5	1.5
Buffer	5 µL	10	10	5
$MgCl_2$	1.25 µL	-	-	3 µL $MgSO_4$
dNTPs (10 mM each)	0.5 µL	1	1	5
Polymerase	0.05 µL	0.5	0.5	1
Glycerol 50% (v/v) (optional)	-	-	-	10 µL
$_{dd}H_2O$ to	25 µL	50 µL	50 µL	50 µL

Table 2.2-2: GoTaq Flexi & GoTaq G2 Flexi PCR cycle protocol.

Stage	**Step**	**Duration**	**Temperature**
A.	Initial denaturation	2 min	95°C
B.	33 cycles		
	1. Denaturation	30 sec	95°C
	2. Annealing	45 sec	Lowest T_m
	3. Elongation	1 min/kb	72°C
C.	Final elongation	10 min	72°C

Table 2.2-3: NEB Taq cycle protocol.

Stage	Step	Duration	Temperature
A.	Initial denaturation	2 min	95°C
B.	33 cycles		
	1. Denaturation	30 sec	95°C
	2. Annealing	30 sec	Lowest T_m
	3. Elongation	1 min/kb	68°C
C.	Final elongation	5 min	68°C

Table 2.2-4: Q5 high fidelity PCR cycle protocol.

Stage	Step	Duration	Temperature
A.	Initial denaturation	30 sec	98°C
B.	25 - 35 cycles		
	1. Denaturation	10 sec	98°C
	2. Annealing	30 sec	Lowest T_m
	3. Elongation	30 sec/kb	72°C
C.	Final elongation	2 min	72°C

Table 2.2-5: KOD hot start PCR cycle protocol.

Stage	Step	Duration	Temperature
A.	Initial denaturation	2 min	95°C
B.	20 - 40 cycles		
	1. Denaturation	20 sec	95°C
	2. Annealing	10 sec	Lowest T_m
	3. Elongation	25 sec/kb	70°C
C.	Final elongation	2 min	70°C

Table 2.2-6: Phusion PCR cycle protocol.

Stage	Step	Duration	Temperature
A.	Initial denaturation	2 min	98°C
B.	30 – 33 cycles		
	4. Denaturation	15 sec	98°C
	5. Annealing	30 sec	Lowest T_m – 5°C
	6. Elongation	30 sec/kb	72°C
C.	Final elongation	10 min	72°C

2.2.1.2 Overlap extension PCR

Overlap extension polymerase chain reaction (overlap extension PCR) was performed as described by Wurch *et al.* 1998 with minor changes.

The genes intended for the creation of gene fusions were amplified from genomic DNA of *E. coli* and *B. subtilis* using Phusion polymerase and primers that had an overlapping region of at least 20 bp for the desired fusion genes. The products of these PCRs were gel purified and 13 µL each were used as template for the overlap extension PCR. The products were gel purified (2.2.2.1, page 68 and 2.2.2.3, page 68) and transformed into *E. coli*.

Table 2.2-7: Overlap extension PCR reaction mix.

Content	Amount
5’ template	13 µL
3’ template	13 µL
5x buffer	10 µL
dNTPs (10 mM each)	1 µL
DMSO	2.5 µL
Phusion polym.	0.5 µL
Ad. $_{dd}H_2O$	50 µL

Table 2.2-8: Overlap extension PCR cycle protocol.

Step	Sub-Step	Duration	Temperature
A.	Initial denaturation	2 Min	98°C
B.	15 cycles		
	1. Denaturation	15 sec	98°C
	2. Annealing	30 sec	55°C
	3. Elongation	30 sec	72°C
C.	20 cycles		
	4. Denaturation	15 sec	98°C
	5. Annealing	30 sec	72°C
	6. Elongation	60 sec	72°C
D.	Final elongation	10 min	72°C

2.2.1.3 pGEM-T Easy

The pGEM-T Easy vector system is convenient for direct cloning of PCR products without the need for restriction enzymes. The pGEM-T Easy vector (Promega) is pre-digested with EcoR V and has single strand 3'-T overhangs on both ends. Poly A overhangs added to PCR products by non-proofreading PCR enzymes thus allow for direct ligation of the PCR product into the pGEM-T Easy vector. The vector carries T7 and SP6 promoters flanking the MCS. The latter is located within the beta-galactosidase gene, allowing for blue/white selection of transformants.

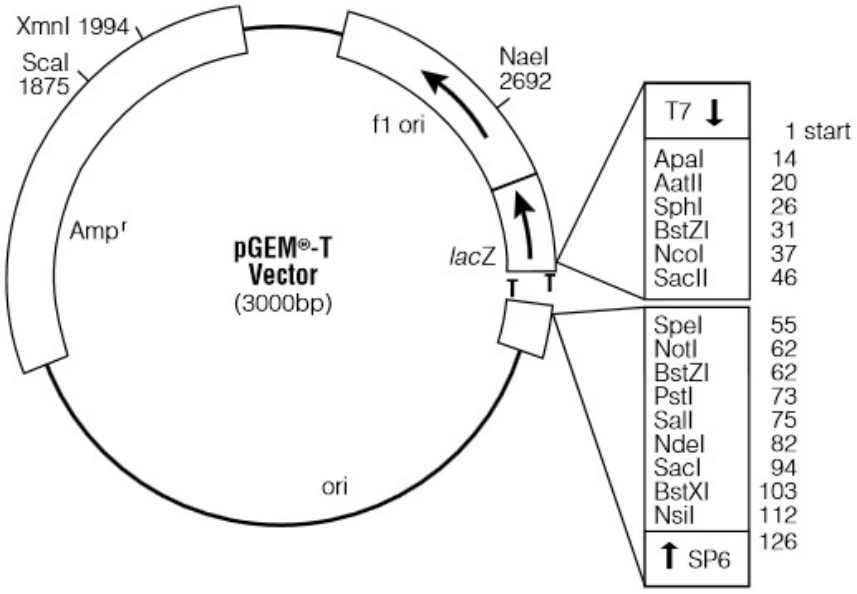

Figure 2.2-1: Vector map of pGEM-T Easy.
From technical bulletin TM042, Promega Corporation 2015.

2.2.1.4 Gibson assembly®

DNA fragments for recombination with the Gibson method (Gibson et al. 2008) were synthesized in separate Phusion polymerase PCRs and contained overlapping regions of about 40 bp. In this project the commercially available master mix for Gibson assembly from NEB was used according to the manufacturers' recommendations.

2.2.1.5 PaperClip assembly

PaperClip is based on the assembly of individually amplified DNA fragments (parts) with the help of specific oligonucleotides (clips) in a PCR based reaction (Trubitsyna

et al. 2014). Parts were amplified using a proof reading DNA polymerase. To minimise template carry over, completed PCR reactions were supplemented with 1 µL of the restriction enzyme DpnI and incubated at 37°C for 1 h. The reactions were loaded on 3% (w/v) agarose gels, bands of desired sizes were cut from the gel and the contained DNA was purified using Qiagen gel extraction kits. Oligonucleotides specific to the beginning and end of each part were annealed and ligated to create clips (figure 2.2-2). The clips were then used to arrange the parts in the desired order during the assembly reaction and allow for amplification of the desired construct (figure 2.2-2 E).

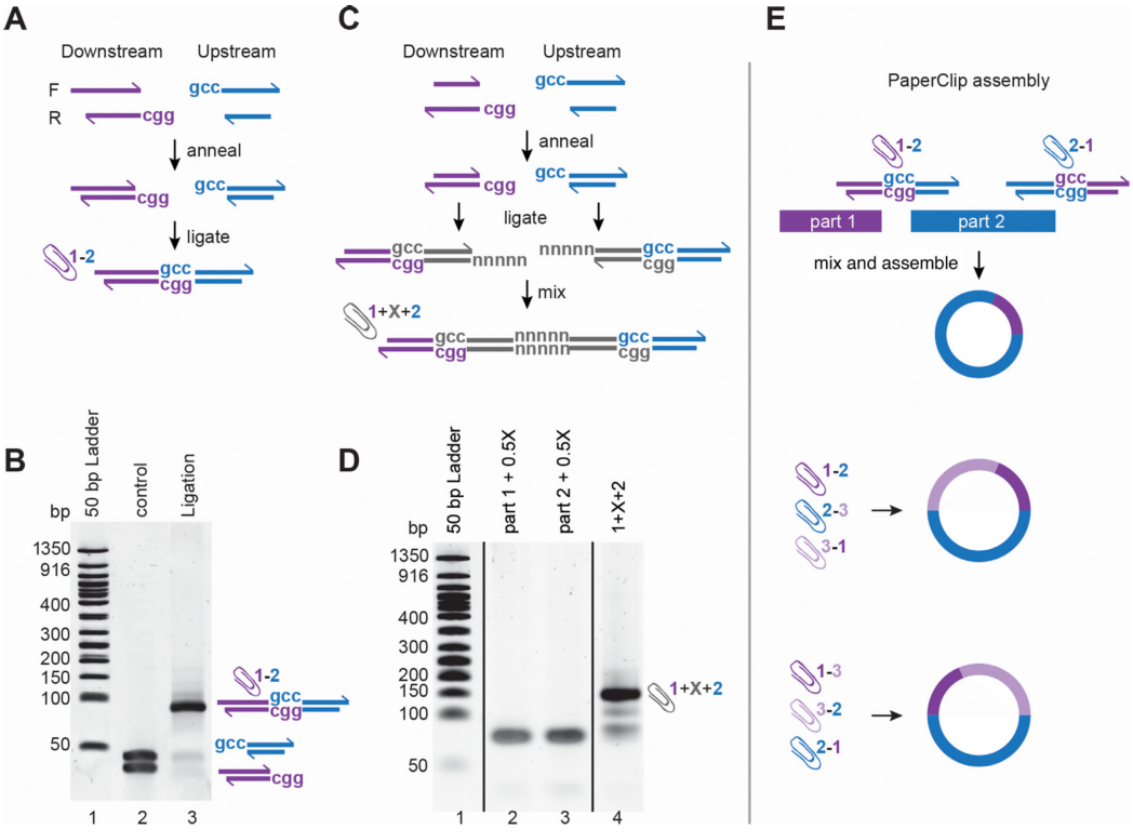

Figure 2.2-2: Method for preparing Clips.

A) Scheme of Clip preparation—oligonucleotides corresponding to the downstream region of one part (in purple) and two for the upstream region of the second part (in blue) are annealed pairwise and ligated. Since the inner parts contain GCC sticky ends and the outer ends non-compatible overhangs, the ligation reaction is correct and efficient (B). C and D) Scheme of addition of an intervening sequence between the Clips. The intervening sequence (in grey) is synthesized in the form of four oligonucleotides, forming GCC and GGC overhangs, which are ligated with the Downstream and Upstream half-Clips respectively. The two reactions are then mixed together and the sequence of interest is restored, guided by five base pre-designed overhangs. The efficiency of ligation is sufficient to perform standard PaperClip assembly. E) Schematic diagram of two part assembly using Clips. For two part assembly two DNA parts and two Clips are mixed together and either PCR or cell extract mediated recombination is performed to assemble a circular plasmid. For assemblies with three or more parts, the order of the parts, 1–2–3 or 1–3–2, is guided by the order of the half-Clips. Since all half-Clips have identical GCC overhangs, any order of the parts can be predefined." (reprinted with permission from Trubitsyna et al. 2014, copyright 2014 Oxford University Press).

For the preparation of half clips, 20 µL of each the forward and reverse oligonucleotide were mixed with 10 µL of ddH_2O and heated to 95°C for 5 minutes in a thermocycler. Reduction of the temperature to 10°C at a ramp rate of 0.1°C/sec allowed the oligonucleotides to anneal and form the double stranded half clip with the GCC or GGC overhang (figure 2.2-2).

Half clips were used to create clips according to table 2.2-9, specific for each combination of parts. The resulting clips were visualised on 2% agarose gels for control of near total ligation.

Table 2.2-9: Clip preparation

Step	**Action**
1	Mix 2.8 µL of the leading "down" half clip and the trailing "up" half clip, add 0.8 µL T4-Ligase buffer and 0.4 µL T4-Polynucleotide Kinase (PNK).
2	Incubate at 37°C for 45 minutes.
3	Add 0.8 µL 10mM ATP and 0.4 µL T4-Ligase.
4	Incubate at 16°C for 90 minutes.
5	Inactivate the enzymes by incubation at 65°C for 20 minutes.

The assembly reaction contained clips and parts in an approximate molar ratio of 20:1 and was carried out using KOD hot start polymerase according to table 2.2-10 and table 2.2-11. Two part assemblies were carried out with a reduced cycle count of five in contrast to the standard cycle count of twenty. Five microliters of the assembly reactions were used for the transformation of *E. coli* competent cells (2.2.3.2, page 69).

Table 2.2-10: PaperClip assembly reaction mixture.

Ingredient	Amount / concentration
Parts	3.3 nM final concentration each
Clips	66 nM final concentration each
KOD buffer 10x	5 µL
dNTPs (2 mM each)	5 µL
$MgSO_4$ (25 mM)	3 µL
Glycerol (50% v/v)	5 µL (for GC-rich sequences only)
KOD hot start polymerase	1 µL
ddH_2O	to 50 µL final volume

Table 2.2-11: PaperClip assembly reaction conditions.

Stage	Step	Duration	Temperature
A.	Initial denaturation	2 min	95°C
B.	5 / 20 cycles		
	1. Denaturation	20 sec	95°C
	2. Annealing & Elongation	30 sec/kb	70°C

2.2.1.6 Circular polymerase extension cloning (CPEC)

Circular polymerase extension cloning (CPEC) is a scar-less PCR based assembly method. In comparison with other PCR based assembly methods the advantages of CPEC are the use of only one enzyme and the general simplicity and robustness of the method (Tian & Quan 2009). However, since an individual primer set is necessary for each assembly, CPEC is comparatively cost and time intensive compared to more modular assembly methods like PaperClip, which only requires one specific primer set for each part which can be combined with any other part in any desired orientation.

For a CPEC assembly, the desired parts are amplified in a PCR reaction, using primers that create 25 bp overhangs which are complimentary to the neighbouring part in the assembled construct. These overhangs are only generated for one side of

the assembly site. This results in double stranded, blunt ended PCR products that, when melted into single stranded DNA, can prime each other in a second PCR reaction via the 25 bp overhang of one side of the assembly site towards the other.

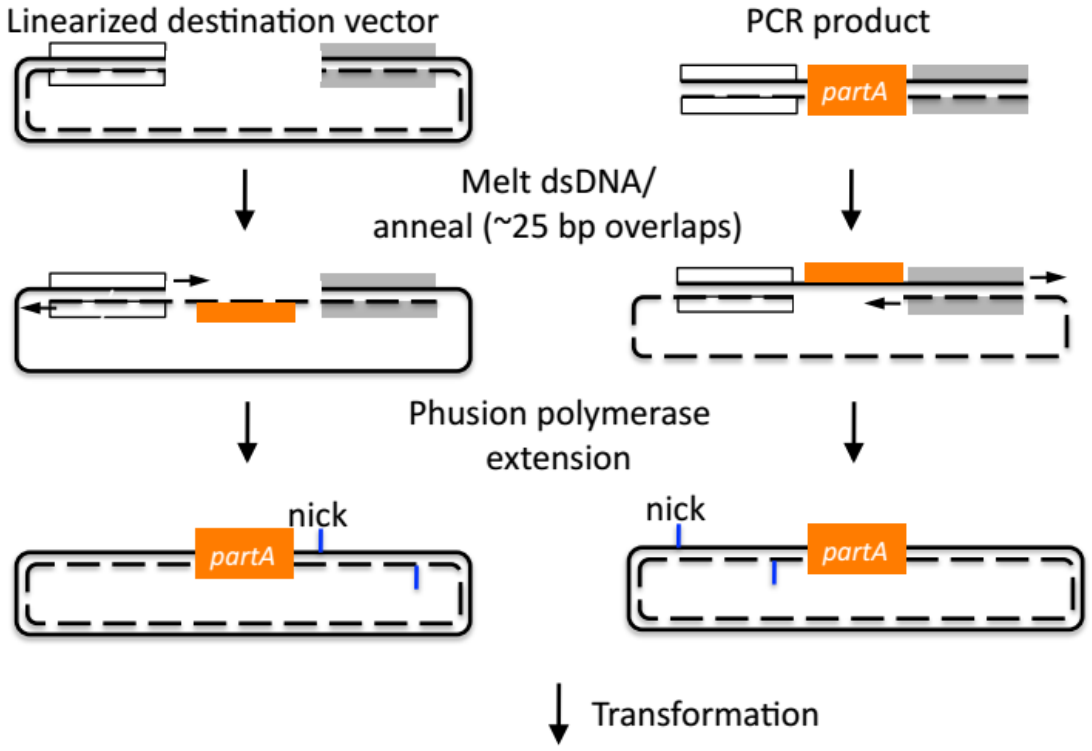

Figure 2.2-3: CPEC assembly schematic (Hillson n.d.).

The destination vector and desired part(s) are amplified in separate PCR reactions, creating sequence identical overhangs of approximately 25 bp. Once molten into single stranded DNA, these molecules can prime each other in a second PCR reaction, resulting in new, circular and double stranded DNA molecules ready for transformation.

Parts for CPEC assembly were amplified in normal PCR reactions using the KOD hot start polymerase (2.2.1.1, page 55), DpnI digested (2.2.1.8, page 67) and run on an agarose gel (2.2.2.1 & 2.2.2.2, page 68). The assembly PCR reaction was composed according to table 2.2-12 and performed following the protocol outlined in table 2.2-13. 5 µL of the final PCR product were used to transform competent *E. coli* cells (2.2.3.2, page 69).

Table 2.2-12: Composition of a CPEC assembly PCR.

Ingredient	Amount / concentration
Linearised vector backbone	100 ng
Other assembly pieces	equimolar to backbone
KOD buffer 10x	2.5 µL
dNTPs (2 mM each)	2.5 µL
$MgSO_4$ (25 mM)	1.5 µL
Glycerol (50% v/v)	2.5 µL (for GC-rich sequences only)
KOD hot start polymerase	0.5 µL
ddH_2O	to 25 µL final volume

Table 2.2-13: CPEC assembly PCR protocol.

Stage	Step	Duration	Temperature
A.	Initial denaturation	2 min	95°C
B.	10 - 15 cycles		
	7. Denaturation	20 sec	95°C
	8. Annealing	10 sec	55°C
	9. Elongation	25 sec/kb	70°C
C.	Final elongation	5 min	70°C

2.2.1.7 Mutagenesis with blunt-end ligation (MABEL)

Mutagenesis with blunt-end ligation (MABEL) is a simple technique widely used to introduce small single locus sequence changes into circular DNA constructs. Two divergent non-overlapping primers (forward and reverse) are used to amplify a circular DNA template in a standard PCR reaction (2.2.1.1, page 55). Desired sequence changes in relation to the template can be introduced by designing a primer with a non-complimentary tail at the 5’ end. To minimise template carry over, completed PCR reactions were supplemented with 1 µL of the restriction enzyme DpnI and incubated at 37°C for 1 h before gel purification (2.2.2.1 & 2.2.2.2, page

68). To ligate the blunt end PCR products, phosphate groups were added to the purified DNA by use of T4 Polynucleotide Kinase (PNK) and ligation was performed with T4 DNA Ligase according to manufacturer recommendations.

2.2.1.8 DNA restriction

DNA was cut with the use of DNA restriction enzymes. All enzymes used were supplied by NEB and used to the manufacturer's recommendation.

2.2.1.9 Ligation

Ligations were performed with an approximate ratio of insert to vector of 1:4 as determined by spectrophotometric measurement of DNA at 260 nm. Reactions were incubated overnight at 16°C or for 1 hour at room temperature, using NEB T4 DNA Ligase.

Table 2.2-14: Standard ligation reaction.

Ingredient	Amount / concentration
Vector DNA	1 part
Insert DNA	4 parts
T4 Ligase Buffer	1 µL
T4 DNA Ligase	0.5 µL
ddH_2O	to 10 µL final volume

2.2.1.10 Removal of phosphate groups of linearized DNA

Self-ligation of linearized vectors and other restriction digested DNA fragments can be inhibited by removal of the phosphate groups from the 5' end of DNA strands. CIP reactions (table 2.2-15) were incubated at 37°C for 30 minutes and column purified (2.2.2.3, page 68).

Table 2.2-15: CIP reaction.

Ingredient	Amount / concentration
Linearized DNA	50 – 250 ng
CutSmart Buffer (10x)	2 µL
CIP	1 µL
ddH_2O	to 20 µL final volume

2.2.2 DNA analysis and purification

2.2.2.1 Agarose gel electrophoresis

Agarose gel electrophoresis was performed according to Johansson 1972 in BioRad gel tanks at voltages between 40 and 150 V depending on expected fragment sizes and desired accuracy. Gels were post stained using GelGreen or pre stained using SaveView according to the manufacturer's recommendations.

Table 2.2-16: TAE-Buffer.

Content	Concentration
Tris base/HCl	40 mM
Glacial acetic acid	20 mM
EDTA	1 mM
pH 8 (NaOH/HCl)	as needed

2.2.2.2 Gel extraction

DNA from Agarose gels was purified with the QIAquick gel extraction kit from QIAGEN according the manufacturers' specifications and eluted in volumes between 15 and 30 µL.

2.2.2.3 Column purification of genomic, plasmid and linearized DNA

Genomic, plasmid and linearized DNA was purified with the QIAquick PCR purification, the QIAprep spin miniprep or the DNeasy blood & tissue kits from

QIAGEN. All purifications were performed according to manufacturer specifications.

2.2.3 Microbiological techniques

2.2.3.1 Microbial cultivation

Unless otherwise stated, overnight cultures of *E. coli* were prepared by inoculating 5 mL lysogeny broth (LB) medium from agarose plates using inoculation loops (Table 2.1-6, page 52). Cultivation took place for 16-20 hours at 37°C in an orbital shaker at 200 rpm under appropriate antibiotic selection. Similarly, cultures on LB-Agar plates were cultivated as aforementioned.

Culture samples that were to be stored for extended periods of time were mixed with 50% (v/v) glycerol in a 7:3 ratio of overnight culture to glycerol and stored at -80°C.

2.2.3.2 Transformation

Competent *E. coli* cells were prepared by a modified method of Chung et al. 1989. Cells of the desired strain were streaked from a -80°C stock onto LB-Agar plates without antibiotic selection and incubated at 37°C overnight. 50 mL LB in a 250 mL Erlenmeyer flask were inoculated from this plate and incubated at 37°C and 200 rpm until an optical density at 600 nm (OD_{600}) of 0.2 to 0.5 was reached. The flask was cooled down quickly in ice water and the culture was centrifuged for 15 minutes at 4,500 rcf at 4°C. After the supernatant was discarded, the cell pellet was resuspended in 5 mL sterile filtered fresh ice-cold transformation and storage solution (TSS, table 2.2-17 below) and transferred into pre-chilled 1.5 µL microcentrifuge tubes in 100 µL aliquots. Aliquots of competent cells were transferred to a -80°C freezer immediately and kept for future use.

For transformations, the appropriate number of aliquoted competent cells were removed from the -80°C storage and placed on ice for several minutes to thaw. Depending on the DNA used for transformation, 1 µL of column purified plasmid DNA (2.2.2.3, page 68) or 5 µL of ligation (2.2.1.6, page 64) or PaperClip assembly reaction (2.2.1.5, page 60) were added to the competent cells and incubated for 30

minutes. The reactions were heat shocked at 42°C for 90 seconds and then transferred back to ice for a further 90 seconds. After the addition of 900 µL of room temperature LB, the cells were incubated at 37°C and 200 rpm shaking for 60 (ampicillin and carbenicillin resistance) to 90 minutes (all other antibiotic resistances). Transformations were then centrifuged at 4,000 rcf for 4 minutes and approximately 800 µL of the supernatant decanted. The cell pellets were resuspended in the remaining supernatant and plated onto LB-Agar plates containing appropriate antibiotics for incubation overnight (2.2.3.1, page 69).

Table 2.2-17: Transformation and Storage Solution (TSS).

Ingredient	**Amount**
Lysogeny broth	17 mL
PEG 3350 (40% w/v)	5 mL
$MgCl_2$ (1 M)	1 mL
DMSO	1 mL

2.2.3.3 Protein expression

Recombinant proteins were expressed in *E. coli* BL21 (table 2.1-2, page 47). Isopropyl β-D-1-thiogalactopyranoside (IPTG) induced expression cultures were grown in Terrific Broth (TB, table 2.1-7, page 53) without antibiotics. Each 50 mL expression culture was inoculated with 500 µL of an overnight culture or glycerol stock, which had been grown in LB with antibiotic selection. Protein expression was induced by adding IPTG to a final concentration of 0.5 mM when cultures had grown to an OD_{600} of approximately 0.6. Samples were taken regularly after induction to monitor for an end of exponential growth.

Autoinduction cultures were grown in autoinduction medium (table 2.1-9, page 54) and routinely started to express recombinant proteins after approximately 16-20 hours incubation at 30°C or 37°C and 300 rpm. Cultivation was stopped after a maximum of 36 hours depending on visible recombinant protein expression and final OD_{600} measurements. The basis of this methodology has been described by Studier in 2005, however the precise protocol used in this work was developed by Trevor Sweeney at the Imperial College, London and has not been published.

2.2.3.4 Ultrasonication

Ultrasonication was performed in Sanyo SoniPrep 150 sonicator at 23 kHz with an amplitude of 15 micron employing an exponential microprobe with a tip diameter of 3 mm. Unless otherwise stated, *E .coli* cells were lysed with 5 s sonication bursts and 25 s cooling period in an ice water bath for 10 cycles.

2.2.4 Protein purification and analysis

2.2.4.1 Protein weight calculations

Protein weight calculation were performed with the online sequence manipulation suite (Stothard 2000) using AA sequences obtained from the translation view of the according open reading frames in the Snap Gene sequence manipulation software (version 3.1.8 and later, GSL Biotech LLC, Chicago, IL, USA).

2.2.4.2 Histidine tag based protein purification via IMAC

Protein purification employing a poly histidine (His) tag on the C- or N-terminal end of an amino acid sequence presents one of the easiest purification methods for most proteins (Lichty et al. 2005). His tag purification was performed with pre packed immobilised metal ion affinity chromatography (IMAC) columns from GE Healthcare (HisTrap HP, 1 mL) according to the procedure outlined in table 2.2-18 below, employing the buffers given in table 2.2-19 to table 2.2-21.

Table 2.2-18: Histidine tag purification protocol.

Step	Action
1	Wash the column with 5 column volumes of distilled water.
2	Equilibrate the column with 5 column volumes of binding buffer at a flow rate of 1 mL/min.
3	Apply pre-treated sample with a flow rate of 1 mL/min.
4	Wash with 9 column volumes of binding buffer at a flow rate of 2 mL/min.
5	Elute with elution buffers of 100 to 1000 mM imidazole, 1 - 5 column volumes each and collect in separate fractions (step gradient).
6	Regenerate column by washing with 5 column volumes of binding buffer or by stripping and regenerating the column according to the manufacturer recommendations. Columns were filled with 20 % Ethanol for long term storage at 4°C.

Table 2.2-19: Histidine tag binding buffer.

Ingredient	Concentration
Tris pH 8	25 mM
NaCl	150 mM
Urea	8 M (only for urea based purifications)

Table 2.2-20: Histidine tag washing buffer.

Ingredient	Concentration
Tris pH 8	25 mM
NaCl	150 mM
Imidazole	40 mM
Urea	8 M (only for urea based purifications)

Table 2.2-21: Histidine tag elution buffer.

Ingredient	Concentration
Tris pH 8	25 mM
NaCl	150 mM
Imidazole	100 mM, 200 mM, 300 mM or 1000 mM
Urea	8 M (only for urea based purifications)

2.2.4.3 Protein concentration measurement

Protein concentrations were measured by an improved Bradford assay (Bradford 1976, Thermo Fisher) using the Pierce Coomassie Brilliant Blue Plus (Bradford) Assay (catalogue number 23236, Thermo Fisher). This assay has an extended linear response curve and is compatible with more detergents than the traditional Bradford assay. The assays were performed according to the manufacturer's recommendations and standard curves were prepared for each experiment.

2.2.4.4 Sodium dodecyl sulphate polyacrylamide gel electrophoresis

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) was performed after Laemmli 1970 and Sambrook et al. 1989. SDS-PAGE gels in this project were either self-cast or obtained from Bio-Rad and run in 1x Tris-glycine running buffer using the Bio-Rad Mini Protean II and III systems. Gel percentages were chosen on a case by case basis according to the requirements of the experiment. If a stacking gel was present, gels were run at 80V until the dye from the loading buffer had reached the border or the resolving gel. At this point the voltage was increased to 180 V and the run was stopped when the dye from the loading buffer had run off the gel. Gels were stained overnight with Coomassie Brilliant Blue brilliant blue or for 15 minutes to 1 hour with InstantBlue Coomassie Brilliant Blue protein stain (Expedeon). Samples were boiled in loading buffer at 1x concentration for 3 minutes directly prior to loading.

Table 2.2-22: 5x Tris-glycine buffer.

Ingredient	Amount
Tris base	15.1 g
Glycine	94 g
SDS	5 g
ddH_2O	to 1 L

Table 2.2-23: 5x SDS-loading buffer in ddH_2O.

Ingredient	Concentration
Tris Cl pH 6.8	250 mM
DTT	500 mM
SDS	10%
Bromophenol blue	0.5%
Glycerol or Ficoll 400	50% / 7.5%

2.2.4.5 Electrophoretic mobility shift assay

Based on the migration of DNA through a gel matrix in an electric field, the electrophoretic mobility shift assay (EMSA) is a method to identify protein-DNA binding events through their effect on the mobility of the DNA-sample (Harvie et al. 2006; Campbell et al. 2007).

Protein samples are incubated with a DNA-fragment which is expected to bind the protein. These samples are then loaded onto a native polyacrylamide gel and separated electrophoretically. As described in 2.2.2.1, the migration speed of DNA fragments through a gel matrix in an electric field depends on the fragment sizes. Proteins with near neutral surface charge, bound to DNA, potentially slow down the DNA samples, leading to a shift in mobility compared to unbound DNA. This shift depends on the size, conformation and charge of the bound proteins. Native polyacrylamide gels were self-cast according to Sambrook et al. or purchased ready to use and run in Tris/Borate/EDTA (TBE) buffer using Bio-Rad Mini-Protean II and III systems (Sambrook *et al.* 1989, Bio-Rad).

After polymerisation, self-cast gels were run in TBE buffer at 100 V for 60 minutes prior to loading samples. This step was omitted for pre-cast gels. Loaded gels were run at 90 V until the dye from the sample loading buffer had reached the lower gel boundary and began to run off.

Table 2.2-24: 5x Tris/Borate/EDTA buffer.

Ingredient	**Amount**
Tris base	54 g
Boric acid	27.5 g
EDTA 0.5 M pH 8.0	20 mL
ddH_2O	to 1 L

Table 2.2-25: Native polyacrylamide gel composition for a 5% gel.

Ingredient	**Amount**
Acrylamide 30% (w/v)	1.66 mL
TBE buffer 5x	2 mL
Ammonium persulfate 10% (w/v)	70 µL
TEMED	3.5 µL
ddH_2O	to 10 mL

Table 2.2-26: 1x EMSA Binding Buffer 1st generation.

Ingredient	**Final concentration**
Tris pH 7.8	20 mM
EDTA	1 mM
DTT	1 mM
Glycerol	0.3 % (v/v)
Spermidine	50 µM

Table 2.2-27: 1x EMSA Binding Buffer 2nd generation modified after Wu & Rosen 1993.

Ingredient	Final concentration
Tris-Cl pH 7.6	10 mM
KCl	80 mM
EDTA	0.2 mM
The below to be added when the buffer is to be used for EMSA but not for PMMA slide assays	
Ficoll-400	2.5 % (w/v)
DTT	0.2 mM
Sonicated salmon sperm	50 µg/mL
BSA	50 µg/mL

Table 2.2-28: Gel green staining.

Step	Procedure
1	GelGreen 10,000x stock reagent was diluted in ddH_2O 3,300 fold to make a 3x staining solution.
2	An appropriate volume of staining solution was poured into a plastic container to cover the entire gel to be stained and the gel was transferred into the staining solution.
3	Staining was performed in the dark for 20 min under careful agitation.
4	Gels were visualised on an Invitrogen SafeView blue light box and pictures were taken with a Sony DSC-RX100 digital still camera.

Table 2.2-29: Coomassie staining.

Step	Procedure
1	Gels were transferred into a plastic container and InstantBlue Coomassie staining solution (Expedeon) was poured carefully until the gels were covered.
2	Staining was performed for 60 minutes under careful agitation.
4	Gels were visualised by scanning on a Xerox WorkCentre 5765.

Table 2.2-30: List of DNA probes employed in EMSAs and PMMA assays.

Probe	Sequence (bases given in subscript are single strand overhangs)
P_{ars} whole sequence	CCAACTCAAAATTCACACCTATTACCTTCCTCTGCACTTACAC ATTCGTTAAGTCATATATGTTTTTGACTTATCCGCTTCGAAAG AGACACTACCTGCAACAATCAGGAGCGCAAT
ArsR-BS	$_{GGTT}$ACTTACACATTCGTTAAGTCATATATGTTTTTGACTTATCC GCTTCG
ArsR-BS-TC	$_{TTTTTTTTTTCCCCCCCCCC}$ACTTACACATTCGTTAAGTCATATATGTTTT TGACTTATCC
P_{zntA} whole sequence	CGGTCACTTCCTGATCGTCCGCTCGCTGTATCTCTGATAAAAC TTGACTCTGGAGTCGACTCCAGAGTGTATCCTTCGGTTAATGA GAAAAAACTTAACCGGAGGATGCC
P_{zntA} competitor	ACTTGACTCTGGAGTCGACTCCAGAGTGTATCCTTCGGTT
P_{zntA}-BS-TC	$_{TTTTTTTTTTCCCCCCCCCC}$ACTTGACTCTGGAGTCGACTCCAGAGTGTA TCCTTCGG

Table 2.2-31: EMSA binding essay.

Step	Procedure
1	Protein was diluted in 15 µL binding buffer and incubated at RT for 10 minutes. In the last iteration of the EMSA protocol the total final protein amount in the 15 µL volume was 200 fmol. Earlier iterations used 50 fmol or 100 fmol respectively.
2	DNA probes in a total volume of 5 µL binding buffer were added and the essay was incubated for a further 30 minutes at RT. In the last iteration of the EMSA protocol the total final DNA amount in the 5 µL volume was 10 fmol, earlier iterations used 15 fmol or 25 fmol respectively.
3	The complete sample was loaded onto a native PAGE gel and gels were run at 90 V until the stain in the loading buffer of the DNA ladder had reached the foot of the gel

2.2.5 Solid surface techniques

2.2.5.1 Carrier material testing

Carrier materials were tested for adhesion of recombinant proteins. 5 µg of protein in 3 µL EMSA (2nd generation) buffer were placed on the materials. Before the protein solutions dried, the materials were immersed into 1x TE buffer and washed under light agitation for 15 min to 3 days at RT.

2.2.5.2 Covalent binding of DNA to PMMA slides

This methods was modified from Sun et al. 2012. PMMA slides (20x70x2 mm) were prepared from a sheet of PLEXIGLAS XT (Evonik Industries AG, Germany). DNA was added in 0.5 µL and later 1 µL spotting buffer and left to air dry. DNA was bound by UV irradiation at 254 nm with a power of 3 mW/cm^2 for 10 min in a Stratalinker 2400 (Stratagene, CA, USA). Slides were washed under agitation in 0.1x standard saline citrate (SSC) with 0.1% (w/v) SDS for 10 min and thoroughly rinsed in deionised water before being allowed to air dry.

Table 2.2-32: Spotting buffer.

Ingredient	Concentration
NaH_2PO_4 pH 8.5	150 mM
Triton X	0.004% (v/v)

Table 2.2-33: 20 x Standard Saline Citrate (SSC).

Ingredient	Concentration
Sodium chloride	3 M
Trisodium citrate	300 mM
Adjust pH to 7.0 with HCl	

2.2.6 Metal sensing techniques

2.2.6.1 *In vitro* assay for AsGard

See section 3.3.10 for a detailed description.

Table 2.2-34: Potassium phosphate buffer (KPB buffer) after Siddiki et al. 2011.

Ingredient	Concentration
Potassium phosphate pH 6.7	50 mM
NaCl	40 mM
optionally	
Sonicated salmon sperm DNA	50 µg/mL
Protein to be bound	20 µg/mL

Table 2.2-35: Protein binding buffer (PBB) after Siddiki et al. 2011.

Ingredient	Concentration
Potassium phosphate pH 7.4	50 mM
NaCl	40 mM
Sonicated salmon sperm DNA	50 µg/mL
Protein to be bound	20 µg/mL

Table 2.2-36: Potassium phosphate – Tween buffer (KP-T) after Siddiki et al. 2011.

Ingredient	Concentration
Potassium phosphate pH 6.0	10 mM
Tween-80	0.05% (w/v)

2.2.6.2 *In vivo* assay for pBest based sensors

Tests for influence of heavy metals on sensor feedback in vivo were carried out in 24, 48 and 96 well plates and micro cuvettes. Cultivation of *E. coli* carrying relevant biosensor plasmids based on the pBest vector where cultivated in shaking (orbital, 700 rpm) multi well plates at 37°C for varying durations. For 24 well plates the

culture volume was 1 mL, for 48 well plates 500 µL and for 96 well plates 200 µL. Absorption measurements were performed in a Spectrostar Nano plate reader (BMG Labtech). For small sample number fluorescence measurements during reporter testing, a minimum of 100 µL per sample was transferred into 70 µL UV-micro cuvettes and measurements were taken with the Modulus Single Tube Multimode Reader (Modulus, Turner Biosystems, now Promega) with a corresponding fluorescence module installed (see table below). Measurements for all metal influence experiments were performed in a BMG Omega microplate reader (BMG Labtech) in black 96-well plates with clear flat bottoms (Greiner bio-one). Due to the long assay times, lids were used to reduce evaporation effects. Control experiments showed minimal effects of the lids on absorbance and fluorescence measurements. To reduce impact further, fluorescence measurements were taken with the bottom optic, eliminating possible influences from condensation or crosstalk in the clear plastic lid. Fluorescence was measured at 355/460 nm and optical density at 600 nm, both measurements were performed every 6 minutes over a period of at least 8 h.

Table 2.2-37: Measurement modules for the Modulus Single Tube Multimode Reader.

Module	**Excitation wavelength**	**Emission wavelengths**	**Used for**
UV	365 nm	410 – 460 nm	*Sce4297*
Blue	460 nm	515 – 570 nm	Spinach2
Green	525 nm	580 – 640 nm	*mCherry*
Red	625 nm	660 – 725 nm	NA
Absorbance	Absorbance at 600 nm (OD_{600})		
Luminescence	Luminescence measurements (Luciferase activity)		

2.2.6.3 *In vitro* assay for pBest based sensors

Cell free reactions were performed according to the manufacturer recommendations (Promega Corporation 2015a) unless otherwise stated. Reaction volumes were reduced to 10 µL. For each reaction 120 fmol plasmid DNA was supplied. CueR and ZntR were supplied in a 1:50 molar ratio corresponding to a total protein amount of 6000 fmol. ArsR was supplied in molar ratios of 1:0, 1:1, 1:10, 1:100 and 1:500.

Heavy metals were supplied in ddH_2O to the final concentrations stated in the corresponding data graphs. MUG was added to a final concentration of 0.2 mM and positive controls under the control of P_{lac} were induced with 0.5 mM IPTG Reactions were pipetted in a master mix wherever possible to reduce pipetting errors in small volumes and dispensed into a 384-well low volume plate. All reactions were performed in triplicate. Fluorescence measurements were performed in 384-well low volume plates in a BMG Labtech SpectroStar Omega plate reader at 355/460 nm at 37°C over 4 h. Plates were covered with clear adhesive film to prevent evaporation of the samples.

2.2.7 Preparation of cell free transcription translation extracts

After Kwon & Jewett 2015. Shaking *E. coli* BL21 (DE3) liquid cultures in 2 x YPTG media without antibiotics were inoculated from an overnight culture in LB and grown at 37°C to an OD_{600} of approximately 2, equivalent to mid-exponential growth phase. Cells were harvested by centrifugation at 4000 x g at 4°C for 15 min. Pellets were washed in cold buffer A 3 times and wet cell mass was weighed. Pellets were flash frozen in liquid nitrogen and stored at -80°C if preparation was not continued immediately. Thawed or fresh cell pellets were resuspended in 1 mL or 2 mL of buffer A per 1 g of wet cell mass and transferred to 1.5 mL tubes in an ice water bath. Cell suspension was sonicated at a frequency of 20 kHz and 20 micron amplitude following sonication time calculation after Kwon & Jewett. Sonicated samples were centrifuged at 12,000 x g for 10 min at 4°C. Supernatants were carefully transferred into fresh, pre-chilled tubes and flash frozen in liquid nitrogen for storage at -80°C. Total protein amount was quantified by Bradford assay. Reactions were performed in reaction buffer as detailed by Kwon & Jewett (2015).

Table 2.2-38: 4x Amino acid mixture after Sun et al. 2013.

Ingredient	**Concentration**
Each of 20 amino acids	6 mM
Dissolve in 0.1x phosphate buffered saline (PBS). Briefly incubate at 37°C if necessary. Use undissolved components (likely cysteine) as suspension.	

Table 2.2-39: Buffer A

Ingredient	Concentration
Tris acetate pH 8.2	10 mM
Magnesium acetate	14 mM
Potassium glutamate	60 mM
DTT	2 mM

Table 2.2-40: Cell free reaction buffer.

Ingredient	Concentration
ATP	1.2 mM
GTP	0.85 mM
UTP	0.85 mM
CTP	0.85 mM
L-5- formyl-5,6,7,8-tetrahyfrofolic acid	34 µg/mL
E. coli tRNA mixture	170 µg/mL
Potassium glutamate	130 mM
Ammonium glutamate	10 mM
Magnesium glutamate	12 mM
Each of 20 amino acids	2 mM
L-Leucine	10 µM
NAD	0.33 mM
Coenzyme-A	0.27 mM
Spermidine	1.5 mM
Putrescine	1 mM
Sodium oxalate	4 mM
Phosphoenolpyruvate	33 mM
Plasmid DNA	Up to 13.3 µg/mL

2.2.8 Spinach2 screening *in vivo*

Spinach screening in *E. coli* was performed following the procedure published by Pothoulakis et al. (2013). Starter cultures were grown in a shaking liquid culture in LB medium with the appropriate antibiotic selection at 37°C overnight. Main cultures in LB were inoculated from starting cultures with a 1% (v/v) inoculum and grown at 37°C shaking. Upon reaching and OD_{600} of 0.4 cultures were induced by addition of IPTG to a final concentration of 5 mM and then grown shaking for further 2 h. For fluorescence measurements 10 µL of culture were diluted into 100 µL LB and DFHBI-1T (Tocris Bioscience) was supplied from a 10 mM stock in DMSO to a final concentration of 200 µM. Subsequently samples were incubated at 37°C for 10 min to allow the DFHBI-1T to penetrate the cells and then placed on ice for 10 min to aid folding of the Spinach2 aptamer which exhibits 100% fluorescence at 20°C but only approximately 50% fluorescence at 37°C due to secondary structure melting (Strack et al. 2013). Following the aforementioned procedure, fluorescence readings were obtained in a FluoStar Omega plate reader (BMG Labtech) employing a 485 nm excitation filter and a 505 nm emission filter. (Pothoulakis et al. 2013).

2.2.9 Spinach2 screening *in vitro* in the S30 expression system

Screening in the commercial S30 expression system (Promega) was performed in a similar matter to the *in vivo* screening (2.2.8). Cell free reactions were prepared as described previously (2.2.6.3) and incubated at 37°C for 4 h. DFHBI-1T was added to a final concentration of 200 µM and samples were subsequently incubated at 37°C for 5 min. To aid folding of the Spinach2 aptamer samples were then transferred to ice and incubated for 10 min before fluorescence was measured in a FluoStar Omega plate reader (BMG Labtech) using a 485 nm excitation filter and a 505 nm emission filter.

2.2.10 Spinach2 screening *in vitro* with the T7 RNA polymerase

Screening in T7 RNA polymerase reactions was performed in a similar matter to the *in vitro* screening in the S30 extracts (2.2.9). T7 RNA polymerase reactions were performed according to the manufacturers' recommendations (

table 2.2-38) and incubated at 37°C for 4 h or 16 h. DFHBI-1T was added to a final concentration of 200 µM and samples were subsequently incubated at 37°C for 5 min. To aid folding of the Spinach2 aptamer samples were then transferred to ice and incubated for 10 min before fluorescence was measured in a FluoStar Omega plate reader (BMG Labtech) using a 485 nm excitation filter and a 505 nm emission filter.

Table 2.2-41: T7 RNA polymerase reaction.

Ingredient	**Concentration**
10x reaction buffer (NEB)	2 µL
NTPs	0.5 mM each
Template DNA	0.2 – 1 µg
RNasin RNase inhibitor (Promega)	0.5 µL (1 U / µL final concentration)
NEB T7 RNA polymerase	2 µL
Add ddH_2O to a final reaction volume of 20 µL	

3 AsGard

3.1 Abstract

The *cis-trans* interaction based sensor 'AsGard' detects As(III) in water by making the release of the regulatory ArsR protein from its DNA binding partner visible.
To facilitate this visualisation, the ArsR protein was fused to mCherry via a flexible linker peptide, similarly to existing approaches in Arsenic biosensor design. The DNA binding partner in form of a short 46 bp oligomer was immobilised on a plastic (PMMA) support (3.3.9). Initial designs lacked the flexible linker peptide and showed no As responsiveness in electro mobility shift assays (3.3.3) or a newly developed Sepharose assay (3.3.4). Later designs showed responsiveness to As(III) at relevant concentrations but not As(V) in EMSAs (3.3.7) and on the PMMA carrier (3.3.10). Initially tested porous carrier materials had led to direct immobilisation of the fusion protein itself (3.3.8).

3.2 Introduction

Arsenic pollution of drinking water represents a major public health thread in many parts of the world (1.1.1). The situation is exacerbated in countries like Bangladesh and Nepal, where the limited infrastructure and decentralised water infrastructure lead to the near uncontrollable exposure of vast parts of the population to contaminated drinking water from tube wells (World Health Organization n.d.). As chemical detection methods available today are insufficient, biosensors seem to present the only solution to this immense sensing need (Jakariya et al. 2007).
Many biosensors for arsenic in drinking water have been described, yet few are simple enough to be used under field conditions and give results quick enough to be practical (Kaur et al. 2015). Originating from this dilemma, the idea of a truly simple to use sensor has been around for several years. Field work in the context of the Edinburgh arsenic sensor revealed that the quantitative response most sensors are striving for, is not necessarily desired by the local population of Nepal and that a very quick and simple "drink or don't drink" response is preferred in some cases

(David Radford, personal communication). A sensor following this principle would need to fulfil some additional requirements to be of practical use:

- Simplicity of use
- Clear feedback at relevant As concentrations
- No necessity for additional devices or technology
- Adequate shelf life under local ambient conditions
- Low price

The sum of these requirements disqualifies a wide range of biosensor concepts. Any sensor with a response time greater than a few minutes is of little daily value for the quick analysis of a water source directly prior to use. This, in combination with legislative regulations for the release of GMOs, makes any whole cell approach based on today's concepts unrealistic for this specific application (1.2.1). Cell free systems have the potential to be fast enough as long as no transcription / translation response is necessary (1.2.2). The ArsR protein and its binding and unbinding of DNA as a function of As(III) concentration appears to be a potential candidate. Various approaches have shown that the ArsR protein can be used for a robust cis-trans sensor based on measurement of the dissociation state of the protein-DNA complex (Kawakami et al. 2010; Siddiki et al. 2011; Siddiki et al. 2012; Kaur et al. 2015; 1.2.4.2.1). However, even the simplest of these sensors still rely on the measurement of fluorescence with the help of fluorimeters or even more complex quantification. This led to the development of the test strip idea. ArsR could be fused to a visible reporter protein. DNA-probes could be immobilized in a set pattern, e.g. a circle or a line on a carrier material. In arsenic free conditions the now visible ArsR protein should co-localize with the DNA. In the presence of As(III), the complex would dissociate and ArsR would be free to move away from the DNA probes. If this setup could be achieved on paper or a similarly porous material, capillary forces might be able to move the protein away from the DNA probes. Contact of such a test strip with water containing As(III) would lead to the dissociation of the fusion protein, visualized by the deviation of the visible colour from the set pattern and giving a clear "don't drink" output. Contact with As(III)-free water would not lead to any change in appearance. The resulting sensor would resemble a mix between a pH test stripe and a pregnancy test. Due to the dissociation being visualised, false-

negative response would appear unlikely. If any contaminant would destroy the complex, the response would be a false-positive. This test strip approach would allow for maximum ease in terms of use, detection and interpretation. Ideally, such a sensor should be storable at high ambient temperatures and humidity, over a time period of several months and be cheap to manufacture. Based on this idea, a set of aims was set for the development of a sensor to be called "AsGard".

3.2.1 Aims of work presented in this chapter

The AsGard project was to be based on the visualisation of the dissociation of ArsR from a DNA probe consisting of a length of the P_{ars} promoter in the presence of As(III). Consequently the following aims were defined:

- Creation of a fusion protein consisting of ArsR from *E. coli* and a visible marker protein
- Demonstration of unchanged DNA binding and unbinding activity of the fusion protein in response to As(III)
- Identification of carrier material candidates capable of DNA binding and providing protein mobility
- Stable immobilisation of DNA probes onto the identified carrier material(s)
- Binding of ArsR-fusions to immobilised DNA probes
- Demonstration of DNA binding and unbinding activity of the fusion protein in response to As(III) within a useful concentration range on the carrier material
- Development and testing of strategies for maximum shelf-life and As(V) reduction

3.3 Results

3.3.1 First generation cloning and protein expression

E. coli arsR (*arsR)* was amplified by PCR from purified genomic DNA. PCR constructs were assessed on an agarose gels and column purified. Primers for this PCR were chosen to remove the stop codon from the gene sequence. Overlap extension PCR was used to fuse the *arsR* gene 3' to *mCherry* (coding for the monomeric red fluorescent protein mCherry) or *EYFP* (coding for the dimeric enhanced yellow fluorescent protein). Both sequences were obtained from laboratory stocks. Primers were designed to add restriction sites for either BspHI or NdeI at the 5' end of *arsR* and for XhoI at the 3' end of *mCherry* / *EYFP* and to remove the stop codon from both gene sequences to allow for a C-terminal His-tag. PCR products were run on an agarose gel for size control, bands of according size were cut and purified. No further changes were made to the sequences of either the *arsR* gene or *mCherry* / *EYFP*. All versions were cloned into pGemT-Easy (2.2.1.3). Clones in *E. coli* JM109(DE3), verified by colony PCR, were obtained for all constructs. Purified plasmids were digested with enzymes according to the introduced restriction sites to allow restriction-ligation sub-cloning into pET28. The pET28 vector caries sequences for poly histidine tags (His-tags) in different open reading frames on either terminus of the multiple cloning side (MCS). This allows for fusion of the His-tag to an expressed protein at either side, depending on the cloning strategy. The His-tag allowed for later purification of the overexpressed proteins via immobilized metal ion affinity chromatography (IMAC). The *lac* operator in pET28 allows for IPTG inducible expression in *E. coli* and the T7 promoter normally leads to high expression levels.

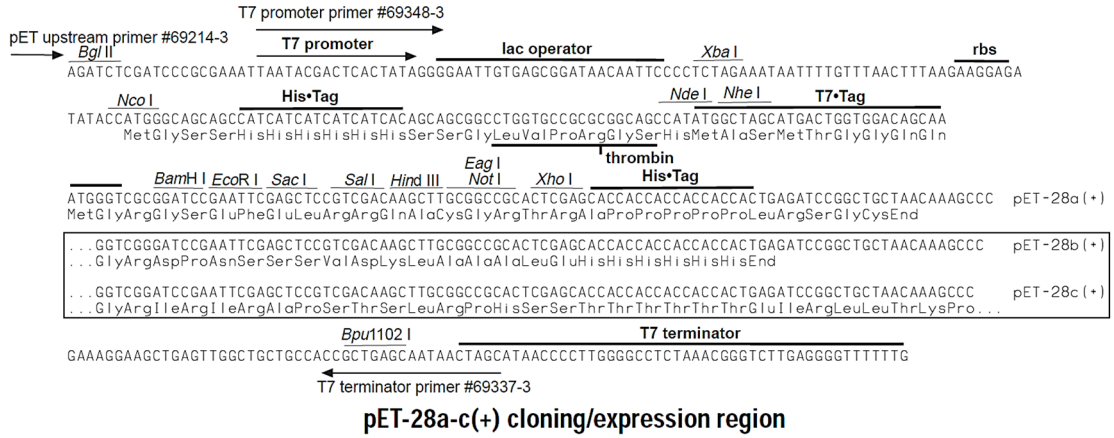

Figure 3.3-1: Cloning / Expression region of pET28 vectors.
Relevant primer binding sites, the T7 promoter and terminator, *lac* operator, RBS, His- and T7-tags as well as restriction enzyme sites are labelled accordingly. pET28 b and c (boxed) sequences are shortened by one or two basepairs respectively compared to pET28a, allowing for adaptation to the open reading frame of the insert. The T7-tag was not used in this work (EMD Millipore).

The pET28 vector was restriction digested with NcoI and XhoI, removing the undesired His-tag for N-terminal fusion, leaving only the C-terminal His-tag. The cohesive end generated by NcoI digest is compatible with the end generated by BspHI digestion. This allowed for ligation without the need for an NcoI digest of the insert DNA. As the *mCherry* sequence contains a restriction site for NcoI, such a digest was undesirable. Following the cloning scheme outlined in figure 3.3-2, two different plasmids were created: pET28-*arsR-mCherry-His* and pET28-*arsR-EYFP-His*. Plasmids were verified by Sanger sequencing and prepared for transformation into the BL21(DE3) strain.

E. coli arsR

```
ATGTCATTTCTGTTACCCATCCAATTGTTCAAAATTCTTGCTGATGAAACCCGTCTGGGCATCGTTTTACTGCTCAGCGAACTGGGAGAG      90
TTATGCGTCTGCGATCTCTGCACTGCTCTCGACCAGTCGCAGCCCAAGATCTCCCGCCACCTGGCATTGCTGCGTGAAAGCGGGCTATTG     180
CTGGACCGCAAGCAAGGTAAGTGGGTTCATTACCGCTTATCACCGCATATTCCAGCATGGGCGGCGAAAATTATTGATGAGGCCTGGCGA     270
TGTGAACAGGAAAAGGTTCAGGCGATTGTCCGCAACCTGGCTCGACAAAACTGTTCCGGGGACAGTAAGAACATTTGCAGTTAA      354
```

mCherry

```
atggtgagcaagggcgaggaggataacatggccatcatcaaggagttcatgcgcttcaaggtgcacatggagggctccgtgaacggccac      90
gagttcgagatcgagggcgagggcgagggccgcccctacgagggcacccagaccgccaagctgaaggtgaccaagggtggcccccctgccc     180
ttcgcctgggacatcctgtcccctcagttcatgtacggctccaaggcctacgtgaagcaccccgccgacatccccgactacttgaagctg     270
tccttccccgagggcttcaagtgggagcgcgtgatgaacttcgaggacggcggcgtggtgaccgtgacccaggactcctccttgcaggac     360
ggcgagttcatctacaaggtgaagctgcgcggcaccaacttcccctccgacggccccgtaatgcagaagaagaccatgggctgggaggcc     450
tcctccgagcggatgtaccccgaggacggcgccctgaagggcgagatcaagcagaggctgaagctgaaggacggcggccactacgacgct     540
gaggtcaagaccacctacaaggccaagaagcccgtgcagctgccccggcgcctacaacgtcaacatcaagttggacatcacctcccacaac     630
gaggactacaccatcgtggaacagtacgaacgcgccgagggccgccactccaccggcggcatggacgagctgtacaagtaataa      714
```

EYFP

```
atggtgagcaagggcgaggagctgttcaccggggtggtgcccatcctggtcgagctggacggcgacgtaaacggccacaagttcagcgtg      90
tccggcgagggcgagggcgatgccacctacggcaagctgaccctgaagttcatctgcaccaccggcaagctgcccgtgccctggcccacc     180
ctcgtgaccaccttcggctacggcctgcaatgcttcgcccgctaccccgaccacatgaagctgcacgacttcttcaagtccgccatgccc     270
gaaggctacgtccaggagcgcaccatcttcttcaaggacgacggcaactacaagacccgcgccgaggtgaagttcgagggcgacaccctg     360
gtgaaccgcatcgagctgaagggcatcgacttcaaggaggacggcaacatcctggggcacaagctggagtacaactacaacagccacaac     450
gtctatatcatggccgacaagcagaagaacggcatcaaggtgaacttcaagatccgccacaacatcgaggacggcagcgtgcagctcgcc     540
gaccactaccagcagaacaccccatcggcgacggccccgtgctgctgcccgacaaccactacctgagctaccagtccgccctgagcaaa     630
gaccccaacgagaagcgcgatcacatggtcctgctggagttcgtgaccgccgccgggatcactctcggcatggacgagctgtacaagtaa     720
taa      723
```

Figure 3.3-2: DNA sequences of *E. coli arsR, mCherry* and *EYFP*.

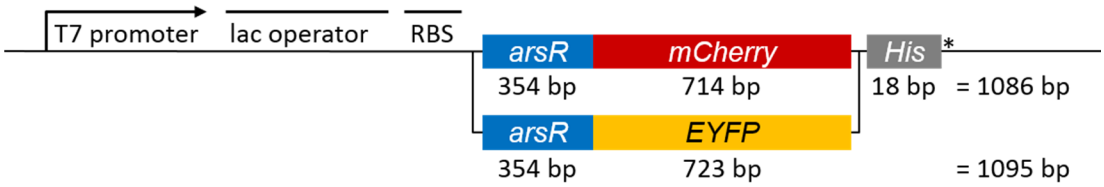

Figure 3.3-3: Schematic cloning scheme in 5'-3' orientation for fusions of *arsR* genes to *mCherry* and *EYFP* in the pET28 vector.

Sequence elements originating from the vector are the T7 promoter, the *lac* operator, the RBS and the His tag as well as the stop codon (*). Relevant sequence lengths are given in bp below the corresponding parts and an approximate total of the sequence length for the fusion proteins is given on the right. The scheme is not to scale.

Transgenic *E. coli* BL21(DE3) colonies were obtained carrying the sequences for the protein fusions ArsR-mCherry-His (ACH) and ArsR-EYFP-His (AEH). ACH was expected to be a 41.03 kDa protein of 361 amino acids (AA) of which 117 AA (13.26 kDa) would form ArsR and 236 AA (26.73 kDa) would form the mCherry protein. The remaining 8 AA would form the His-tag and a 2 AA scar at the end of the mCherry sequence from the digestion-ligation cloning into pET28. The only expected difference of the AEH fusion protein was the slightly larger EYFP at 239 AA (26.98 kDa) making it a total of 41.28 kDa (364 AA). figure 3.3-3 shows a

homology model of the expected fusion proteins. The helix seen in this protein structure, spanning horizontally through the figure, is the α5 helix of the ArsR protein which is thought to be involved in the dimerization of the protein.

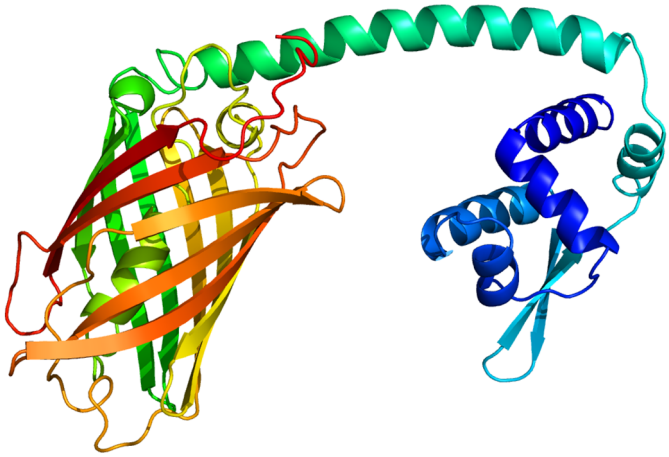

Figure 3.3-4: Possible tertiary protein structure of ArsR-mCherry.
Colours change from deep red for the C-terminus over yellow and green to deep blue for the N-terminus. Accordingly, shown in shades of blue and turquoise, is the anticipated structure for ArsR. The structure prediction is based on homology modelling by the Phyre2 server (Kelley & Sternberg 2009).

Both clones were verified by DNA sequencing. After verification, liquid cultures were inoculated to assess growth rates and induction points in Terrific Broth at 37°C and 200 rpm shaking. Samples were taken every 30 minutes after an initial growth phase of 2 hours. It was found that test cultures entered the exponential growth phase after approximately 4 hours and consequently protein expression was induced after 4.5 hours at an OD_{600} of 0.87 to 0.96. The test cultures were grown for a total of 8.5 hours equalling 4 hours after induction with the OD_{600} reaching maximum values of 2.30 to 3.19 (figure 3.3-4).

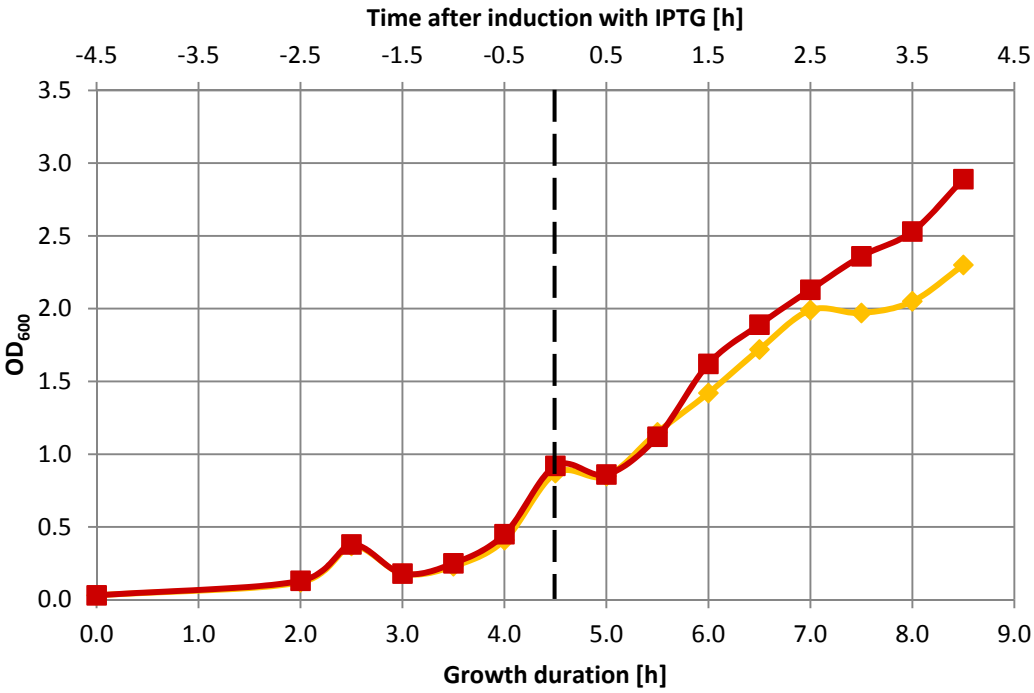

Figure 3.3-5: Growth curves based on OD_{600} measurements.
The time point of induction with IPTG is marked with a dashed vertical line. Timescales are given both in hours after inoculation (bottom) and after induction (top). ■, red: pET28-*ArsR-mCherry-His*; , orange: pET28-*ArsR-EYFP-His*. Cultures were grown at 37°C shaking at 200 rpm in Terrific Broth without antibiotic selection.

The culture expressing ACH showed a visible pale red colour 2.5 h after induction with an increase in intensity up to 3.5 h after induction (figure 3.3-5). The cell lysates of samples taken after the induction were analysed using SDS-PAGE. Clear bands were obtained at the desired protein weight around 41 kDa and the intensity of the bands levelled off after 3 to 3.5 hours of growth (figure 3.3-6 and figure 3.3-7).

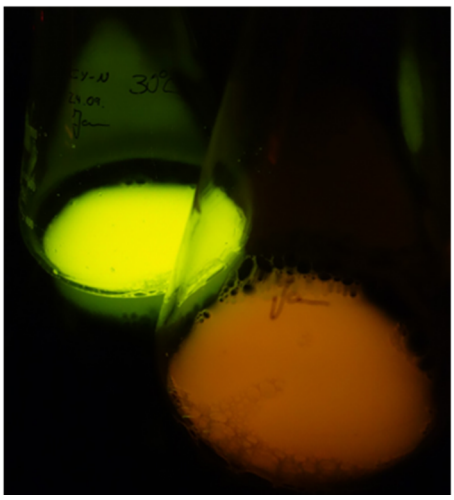

Figure 3.3-6: Fluorescence of fusion proteins ACH and AEH in liquid cultures 4 h after IPTG induction at 37°C.

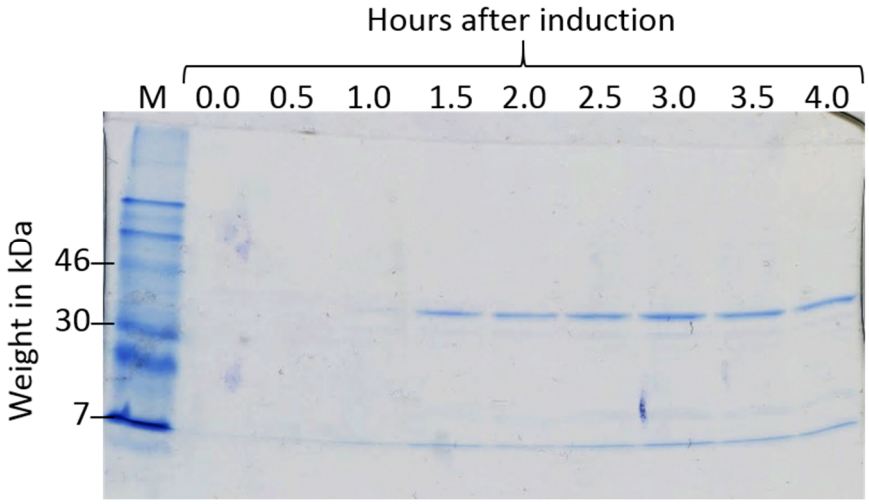

Figure 3.3-7: SDS-PAGE for ACH.

SDS-PAGE stained with Coomassie Brilliant Blue Brilliant Blue showing cell lysates from samples taken every 30 min after induction of the test culture with IPTG. The 7 kDa, 30 kDa and 46 kDa bands of the marker (M) are labelled for easier interpretation. 0.0 – 4.0: time point after the induction of the culture with IPTG in hours.

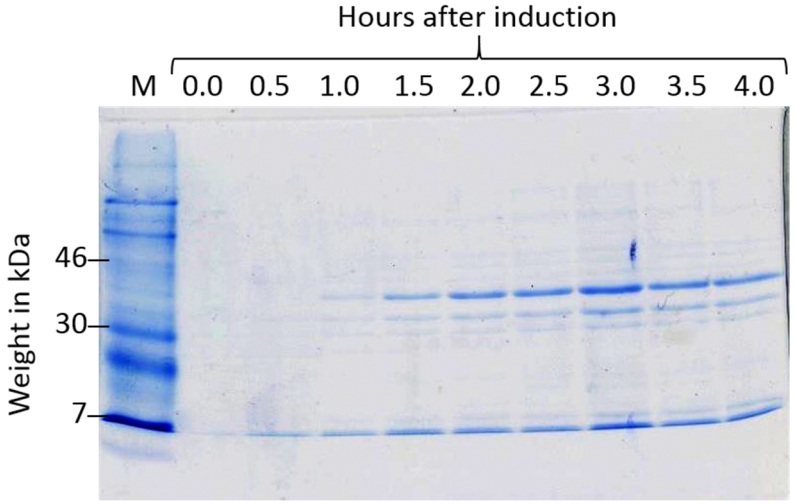

Figure 3.3-8: SDS-PAGE for AEH.

SDS-PAGE stained with Coomassie Brilliant Blue Brilliant Blue showing cell lysates from samples taken every 30 min after induction of the test culture with IPTG. The 7 kDa, 30 kDa and 46 kDa bands of the marker (M) are indicated for easier interpretation. 0.0 – 4.0: time point after the induction of the culture with IPTG in hours.

Based on the data obtained from the test cultures, expression cultures in Terrific Broth were inoculated with a fresh 5 mL overnight culture and induced 3.5 h later upon reaching an OD_{600} of 0.6 to 0.76. The cells were harvested by centrifugation 3.5 h after induction. In later experiments the IPTG induction was replaced by an auto-induction method, partially developed at Imperial College, London (2.2.3.3). Results obtained were identical in regard to increase in protein contents of desired sizes. However, visible fluorescence of the fusion proteins was drastically increased (figure 3.3-8) and final OD_{600} values of around 5.5 were reached. To ensure sufficient aeration, flasks were not filled to more than 10% of their nominal volume and shaking speed was increased to 300 rpm. Cultures were grown at 30°C and 37°C. Since no yellow fluorescence was observed in cultures grown at 37°C and red fluorescence was more intense in cultures grown at 30°C, all subsequent cultures were grown at 30°C. Unless otherwise stated, all further results presented were obtained from lysates and purifications based on auto-induction cultures grown at 30°C.

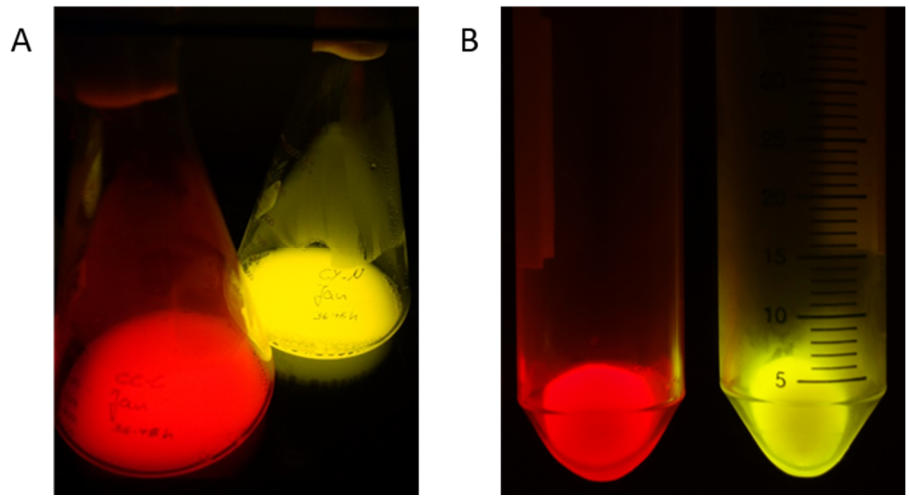

Figure 3.3-9: Fluorescence of liquid auto-induction cultures grown at 30°C expressing ACH and AEH.
ACH fluorescence is visible in bright red in liquid cultures (A) and in cell pellets of the same culture (B). EYFP fluorescence from AEH is visible as bright yellow under both conditions.

3.3.1.1 Protein purification and quantification

The harvested cells were lysed by ultrasonication (2.2.3.4) and the cell lysates obtained were purified via IMAC (2.2.4.2) employing an imidazole step gradient for elution. For both fusion proteins, purifications were performed with and without 8 M urea in the buffers and protein concentrations were measured by Bradford assay. Usage of 8 M Urea was expected to denature most proteins in the lysate and during the purification process, thus ensuring exposure of the His-tag and acting as a control in case of very low protein concentrations being obtained under native purification conditions. The overall supernatant protein yields after sonication of cell pellets derived from 50 mL liquid cultures ranged from 5147 to 6224 µg. Apart from the samples taken for Bradford measurement, the entire supernatant was used for the purification. Yields for purifications with 8 M urea were generally higher than for purifications under native conditions. This difference was more pronounced for AEH with concentrations in the step gradient elution fractions increasing approximately 10-fold in presence of Urea and a reduced protein concentration in the wash fraction. ACH showed comparatively little difference between the two methods, with concentrations increasing less than 2-fold.

Overall concentrations of ACH were significantly higher (over 35-fold) in the 0.2 M imidazole fractions when compared to the corresponding fractions of the EYFP fusion.

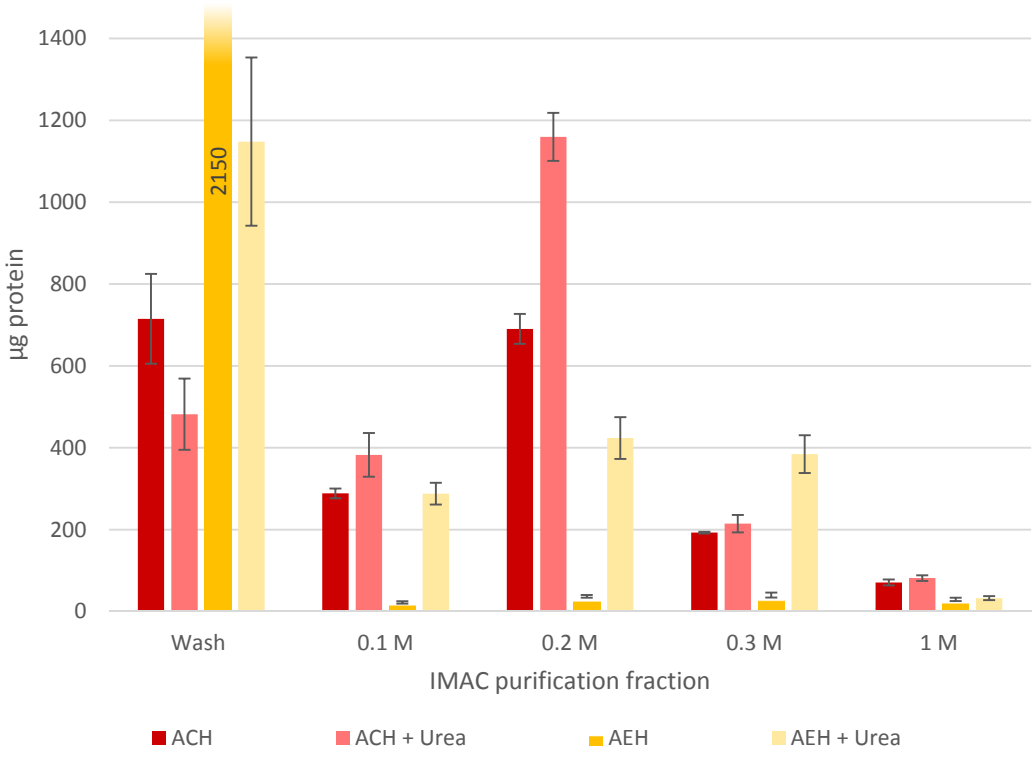

Figure 3.3-10: Total protein amounts before and after IMAC His-purification.
Shown in dark red and light red, ACH and ACH in presence of 8 M urea respectively. Shown in intense yellow and light yellow, AEH and AEH in presence of 8 M urea respectively. The protein concentration in each fraction was measured by Bradford assay and total protein amounts were calculated from these measurements. The assay was performed in triplicate for each sample and error bars represent one SD above and below average. The protein amount in the wash fraction of AEH exceeds the axis maximum chosen for best visibility. The numerical value of the average for this sample is printed into the data bar. SD for this sample was 124 µg.

Matching the results obtained by Bradford assay and SDS-PAGE, most samples from the ACH line, cultures, lysates, pellets and the elution fraction showed a pink colour under daylight, this colour being most intense for samples eluted with imidazole concentrations of 0.1 M or 0.2 M and 8 M urea. figure 3.3-10 gives an impression of the various colour intensities observed.

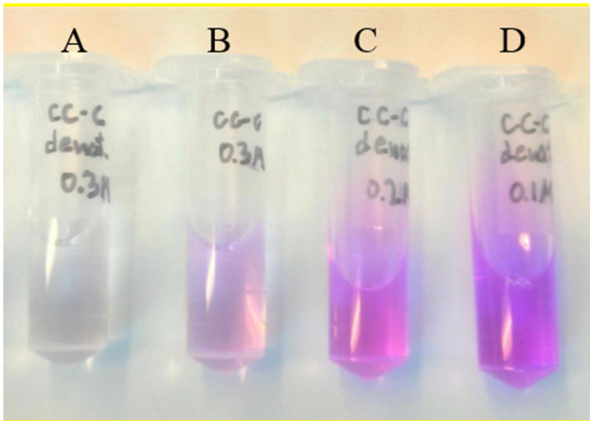

Figure 3.3-11: Visibility of mCherry in various IMAC fractions under daylight conditions. ACH protein fusions eluted at: A, 0.3 M imidazole and 8 M urea; B, 0.3 M imidazole; C, 0.2 M imidazole and 8 M urea; D, 0.1 M imidazole and 8 M urea.

Neither the cultures, nor the protein samples for the EYFP fusions showed any visible colour under daylight conditions. However, some fluorescence under blue light (Safe Imager, Invitrogen, single-intensity signal at 470 nm, bandpass approx. 40 nm) could be observed for samples purified under non-denaturing conditions as well as for the culture and pellet prior to purification. This fluorescence was weak compared to the fluorescence of ACH. The fluorescence intensity for ACH samples eluted in the presence of 8 M urea was brighter than the fluorescence of samples eluted under non-denaturing conditions and the most intense fluorescence was observed at lower imidazole concentrations when compared to non-denaturing conditions (figure 3.3-11). Protein samples containing urea were not used in any further experiments.

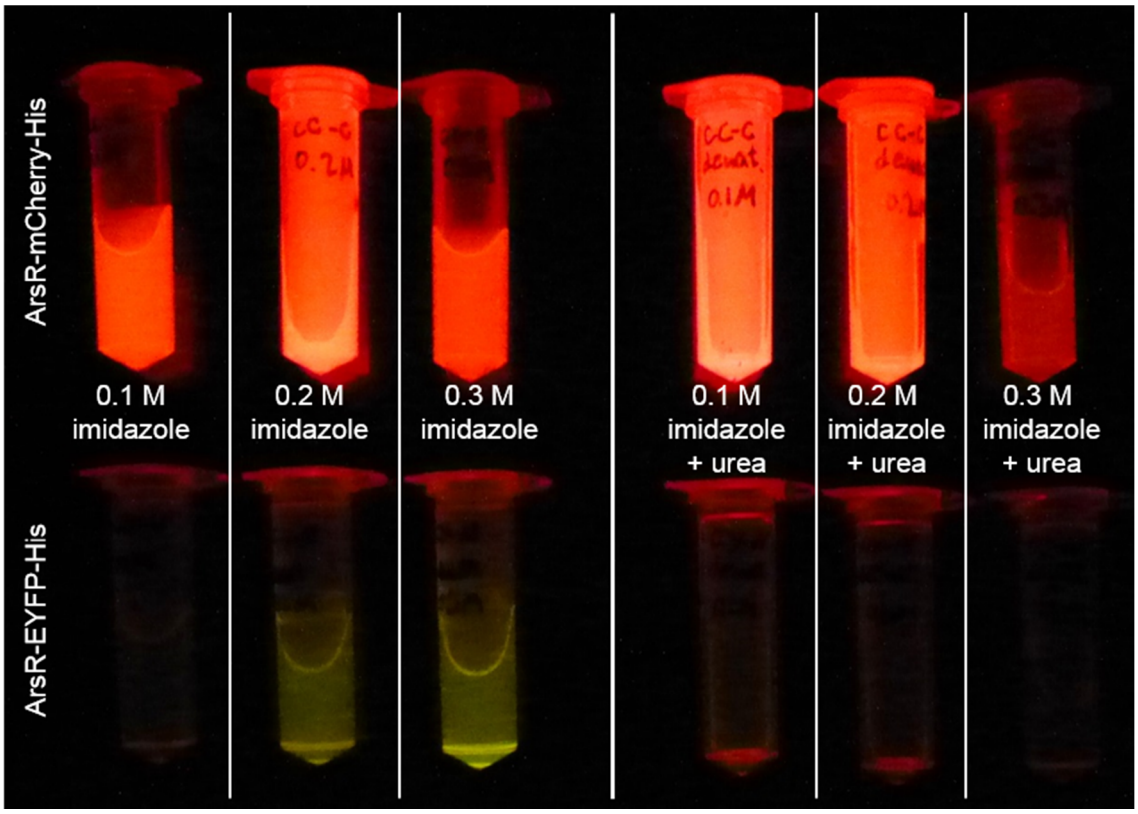

Figure 3.3-12: Fluorescence under blue light (470 nm) of fusion proteins ACH and AEH after purification by IMAC.
The first three columns from the left show protein fusions eluted at 0.1 M, 0.2 M and 0.3 M imidazole respectively, the right three columns show proteins eluted at the same imidazole concentrations but in presence of 8 M urea.

Based on the very limited visibility of EYFP under daylight conditions, the reduced fluorescence under blue light, potential stability issues and the strong performance of mCherry fusion proteins, further experiments were carried out with ACH alone.

3.3.2 Establishment of a non-radioactive electrophoretic mobility shift assay

Based on the results from the protein purification, electrophoretic mobility shift assays (EMSA) were performed for ACH protein fusions eluted with 0.1 M or 0.2 M imidazole (without urea). EMSA is a widely applied technique for the visualisation of protein-DNA binding events and has been used in the past to show the specific binding of DNA probes consisting of various lengths of P_{ars} through ArsR (Xu et al.

1996; Siddiki et al. 2011). Initially the entire 118 bp promoter P_{ars} (2.1.2) was amplified from genomic DNA (*E. coli*) via PCR using proof reading enzymes. PCR products were analysed by means of agarose gel electrophoresis and column purified. The EMSA was initially performed following a protocol established by Campbell at the University of Cambridge in 2007 for the SmtB protein. This protocol uses a low concentration Tris-EDTA binding-buffer at a pH of 7.8, supplemented with DTT (1mM), glycerol (0.3%) and spermidine (50 μM). DTT was left out of samples incubated with metals. Protein samples were equilibrated in this buffer for 10 minutes at RT before being mixed with DNA and incubated for a further 30 minutes at RT before being resolved by PAGE (2.2.4.5).

To remove the need for radioactive labelling and the increasing administrative effort associated with it, a non-radioactive imaging approach had to be developed. PAGE gels were screened under three different conditions: firstly, they were visualised without any staining under blue light, showing the fluorescence of the fusion proteins; secondly, the same gels were examined by staining with Gel Green and subsequent screening under blue light. Gel green is a DNA stain widely used for nucleic acid electrophoresis visualisation in both agarose and polyacrylamide gel electrophoresis. The third and last method was staining with Coomassie Brilliant Blue to visualise any non-fluorescent protein components. Comparison of the three images obtained allowed for clear visualisation of band-shifts (figure 3.3-14).

For initial tests it was deemed acceptable to omit the inclusion of competing proteins and DNA, traditionally added to EMSAs in the form of bovine serum albumin (BSA) and salmon sperm DNA. These were added in later experiments. The Tris-Borate-EDTA PAGE (TBE-PAGE) gels were cast at an acrylamide concentration of 5%. As shown in figure 3.3-12 for the Gel Green stain, initial results showed a need for substantial enhancements in the protocol.

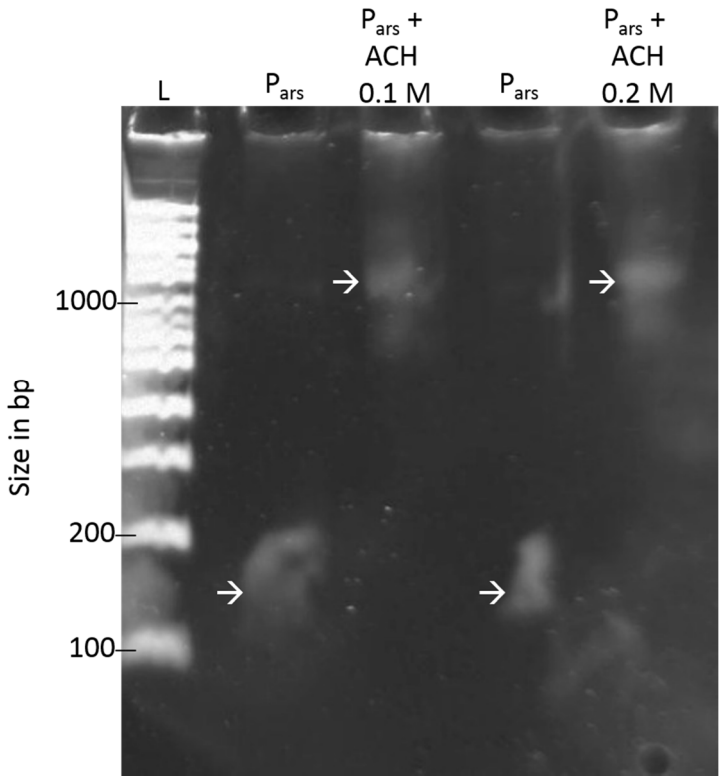

Figure 3.3-13: First EMSAs required significant improvement of the methodology.
Shown is a mobility shift essay of ACH with P_{ars} as specific DNA probe. Bands were believed to be seen at the positions marked by an arrow (→); around the sizes on the DNA ladder (L) indicated on the left. 0.1 M: eluted at 0.1M imidazole; 0.2 M: eluted at 0.2 M imidazole.

The original protocol lacked a dye in the samples, making loading the gel particularly challenging. To minimize gel loading errors, thus enhancing picture quality, a stock of highly concentrated bromophenol blue in $_{dd}H^2O$ was prepared and enough to colour the samples pale blue was added immediately prior to loading the gel, in order to minimise the effect during equilibration and DNA binding. While loading the now visible samples, it became apparent that the amount of glycerol in the buffer was not sufficient to securely keep the samples from rising out of the gel wells. In consequence, the glycerol concentration in the initial EMSA buffer was increased from 0.3% to 1.3%. In addition to P_{ars} as a DNA probe, some protein samples were incubated with a purified DpnI digest of genomic DNA from *E. coli* to show specific binding of P_{ars} by ACH. However, in essays prepared with the modified buffer, not even the weakest of shifting bands was observed (gels not shown). As a consequence, subsequent gels were loaded with samples stained with

Orange G, a dye widely used in pre-coloured buffers for PCR, deemed to have minimal effect on protein-DNA interactions. This partially prevented the effects described previously. However, overall imaging quality remained somewhat unsatisfying (gels not shown). It was decided to change the EMSA buffer to reflect the buffer used by Xu et al. in their publications about ArsR proteins (Xu et al. 1996). The revised buffer was still based on low concentration Tris-EDTA at a slightly lower pH of 7.6 but contained 80 mM KCl and a reduced DTT concentration of 0.2 mM in samples free from metals. Spermidine was omitted and stepwise optimisation of the protocol led to greatly improved assay quality. It was found that the use of 2.5% (w/v) Ficoll-400 (Sigma Aldrich) instead of glycerol increased crispness of bands obtained. Increasing the polyacrylamide concentration to 8% was found to lead to more reliable polymerisation, greatly enhancing separation quality. Later gels were purchased precast (Bio-Rad) and proved to deliver great improvements in EMSA quality. BSA and salmon sperm DNA (both 50 µg/mL) were added to later assays to show specific binding and were found to have no effects on shifting behaviour. BSA had no detrimental effects on visualisation while the salmon sperm DNA led to a slightly increased background in the DNA stainings, as would be expected.

3.3.3 ArsR-mCherry-His shows no metal responsive properties in EMSA

Early mobility shift experiments employing the entire P_{ars} sequence as a DNA probe indicated selective binding of the probe by ACH. The metalloregulated promoter of *zntA* (P_{zntA}) was chosen as a competitive DNA probe. P_{zntA} has some sequence similarities to P_{ars} in containing an inverted sequence repeat in the operator region of the promoter. As shown in figure 3.3-13, addition of P_{zntA} to ACH did not lead to a significant change in the migration pattern on the gel. In contrast, addition of P_{ars} led to the formation of a DNA smear, indicating a DNA-protein interaction event. However, overall quality of the imaged gels remained unsatisfactory and clear shifting bands were elusive.

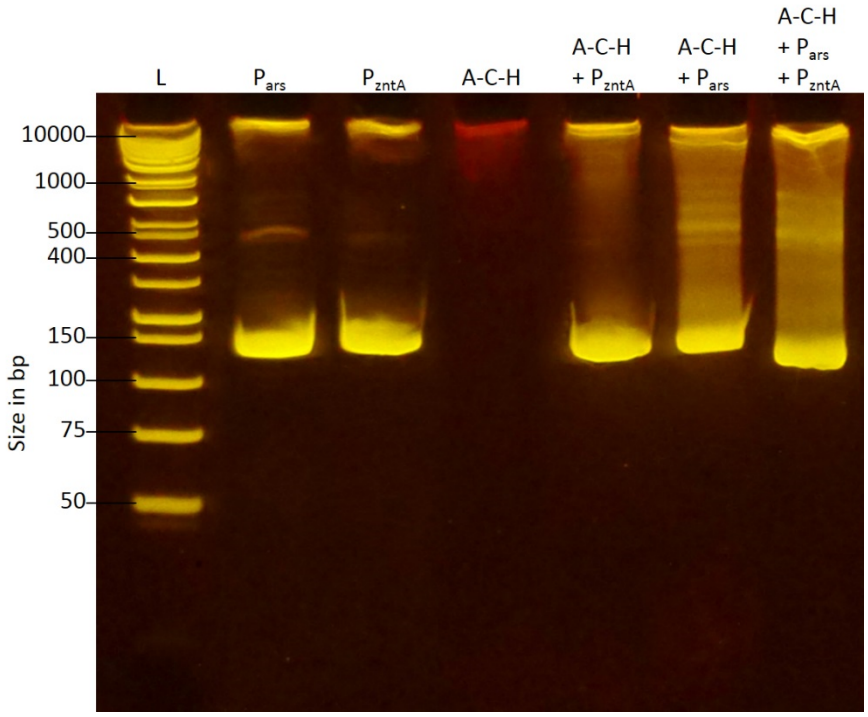

Figure 3.3-14: EMSA of ACH in presence of P_{ars} and P_{zntA}.
Relevant sizes of DNA ladder bands are given in bp. DNA was visualised with Gel Green and screened under blue light. Some red fluorescence from ACH is visible. ACH did not affect the movement of P_{zntA} DNA through the gel in contrast to P_{ars} (→). Approximately 25 fmol DNA probe and 50 fmol of ACH were mixed and pre incubated in 1x EMSA-loading buffer before loading into the appropriate gel wells.

With the aims of increasing the potential shift in protein migration induced by DNA binding and enhancing image quality, the experiment was repeated with shorter DNA probes and reduced overall DNA amounts. A new specific DNA probe, consisting of 47 bp of the operator region of P_{ars} after Xu et al. and a 4 bp, single stranded overhang at either end, was designed (Xu et al. 1996). The idea behind this overhang was that the probe could either be ligated into longer, multi-binding site molecules or connected with other molecules, containing desirable features, such as functionalisations for solid surface binding or imaging techniques (figure 3.3-14).
Competitor probes consisted of 40 bp of the operator region of P_{zntA} incorporating the ZntA binding region (see materials and methods for a full list of all probes used in the various iterations of EMSAs performed).

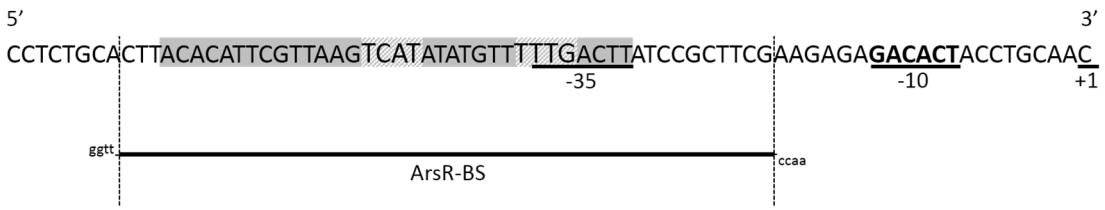

Figure 3.3-15: Partial P_{ars} sequence showing the position of the ArsR-BS DNA probe.
ArsR contact points (Xu et al. 1996) have a hatched background. The shaded sequence including hatched areas indicates the ArsR binding region (Xu et al. 1996) as defined by DNase I foot printing. The -35 and -10 promoter regions as well as the transcription start (+1) are underlined. Single strand overhangs are shown at their respective positions at the end of the probe.

Due to the changed probes, the shift increased and image quality improved. However, no change of DNA-protein binding events was observed upon addition of As(III) in the form of sodium metaarsenite ($NaAsO_2$) to the loading buffer prior to sample incubation at concentrations up to 667.36 µM (equivalent 50,000 ppb As, figure 3.3-15).

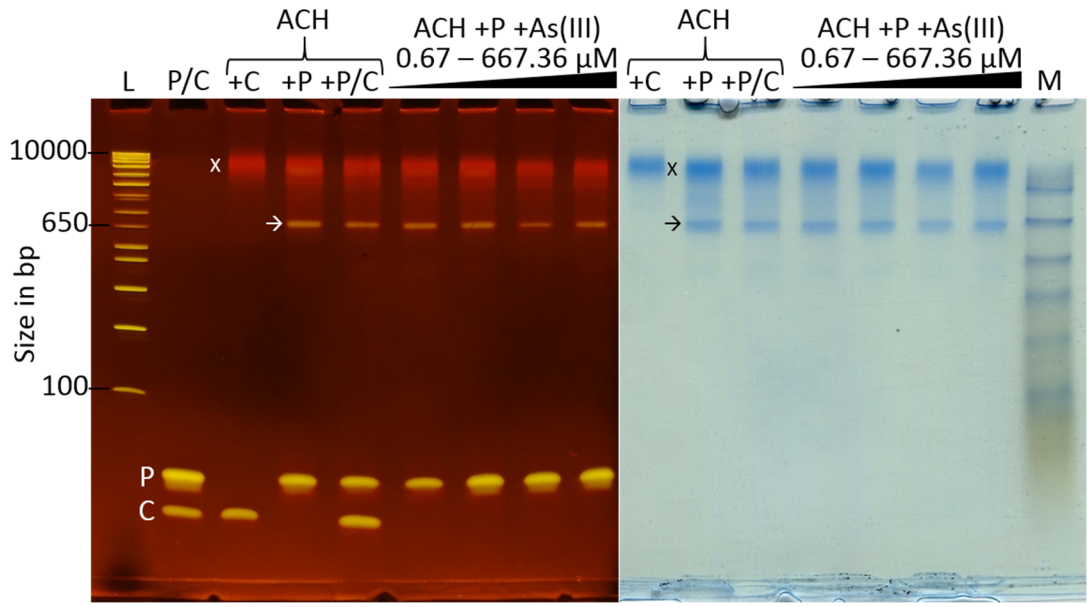

Figure 3.3-16: EMSA of ACH and shorter DNA probes.

Two stains of the same gel are shown, Gel Green staining on the left and Coomassie Brilliant Blue staining on the right. Gel lanes not containing any samples visible with either of the stains have been cut from the images accordingly. Size in bp of relevant bands of the DNA-ladder (L) are indicated. Note that the native protein samples did migrate through the gel as they would in SDS-PAGE and that protein standard labels have been left out for that reason. The position of DNA bands of P_{zntA} competitor probes (C) and specific ArsR-BS DNA probes (P) consisting of ArsR-BS are labelled. DNA-protein band positions are indicated (→) and red fluorescence of ACH is visible in gel areas coinciding with protein presence revealed by Coomassie Brilliant Blue staining (X). Sodium metaarsenate (As(III)) concentrations in the loading buffer used were 0.67 µM, 6.67 µM, 66.74 µM and 667.36 µM. Approximately 10 fmol of each DNA probe, 100 fmol of ACH and the corresponding amount of As(III) were mixed and pre incubated in 1x EMSA-loading buffer before loading into the appropriate gel wells.

No non-specific DNA binding to the P_{zntA} competitor probes (40 bp binding region of P_{zntA}) was observed. All samples containing protein showed fluorescent bands near the gel loading pockets, indicating very limited travel of unbound ACH through the gel during electrophoresis, marked with "x" in the figure above. The protein marker consisting of denatured proteins that had been loaded as a control for denatured protein migration, exhibited normal migration in the native PAGE (lane "M" in the figure above). No unspecific DNA binding to the competitor probes was observed. Coomassie Brilliant Blue staining and fluorescence of mCherry revealed

that a relatively large fraction of the loaded ACH protein did not bind the abundantly present DNA. Variation of the DNA-protein ratio on additional gels did not change this relatively low binding efficiency. Reduction of the protein amounts led to a reduction of the shifted, as well as the unshifted protein bands (not shown). Since this made gels more difficult to image, protein concentrations remained largely unchanged in the following experiments.

Protein purification is a major cost factor in the industrial manufacturing of biosensors. Additionally, multiple buffer changes and physical strain during the purification can have detrimental effects on protein activities and lead to atypical behaviour. Due to this, crude cell extracts were prepared following the same procedure employed for the protein purification via IMAC. After sonication, samples were centrifuged and supernatants were used for an EMSA. Care was taken to keep samples on ice at all times prior to incubation in the binding buffer to reduce the risk of protease induced degradation of the lysates.

Addition of lysates to lanes containing specific DNA probes led to the formation of a putative DNA-protein band at the position previously observed in EMSAs. No such band-formation could be seen upon addition of P_{zntA} competitor DNA probes alone, indicating a specific binding activity. At first, there was concern that the high protein and DNA concentration in the lysates might overstrain the non-radioactive imaging strategy but this was found to be less problematic than anticipated. Especially Gel Green stains showed clear formation of a band at the anticipated position over the background of DNA from the lysate (data not shown). Spontaneous re-binding of ACH and the specific DNA probe after gel loading due to quick dispersion of the As(III) from the loading buffer and ACH metal binding sites into the gel and running buffer was thought to be a possible cause of the unresponsiveness of ACH to As(III) in previous EMSA experiments. As a consequence 667.4 µM As(III) was included in the running buffer and EMSAs were repeated as previously described. Lysate samples were included in this run to preclude a loss of metal binding activity as a consequence of protein purification. Results were clear in that the band pattern was identical to that previously observed, indicating no As(III) responsiveness of ACH (figure 3.3-16).

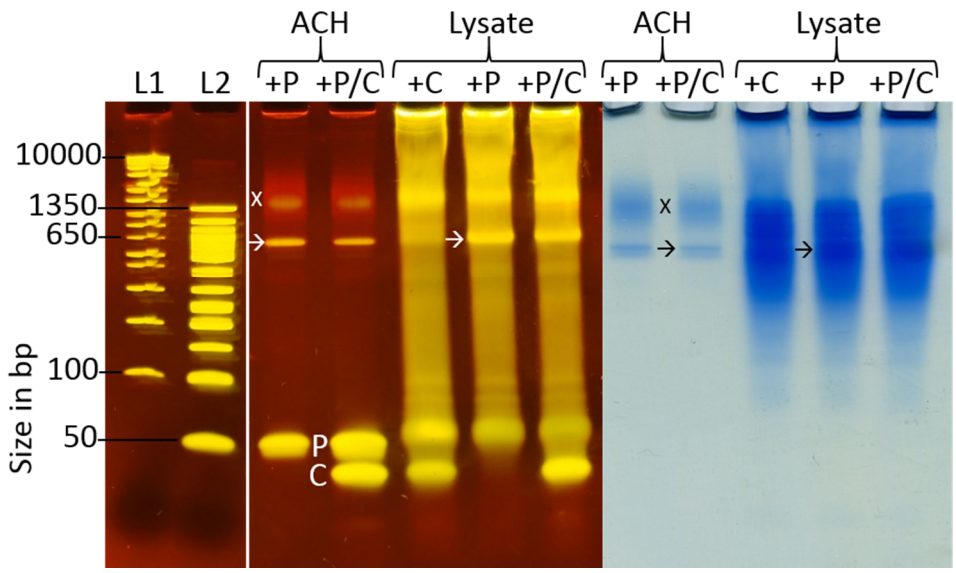

Figure 3.3-17: EMSA of purified ArsR-mCherry-His (ACH) and crude cell lysate.
Two stains of the same gel are shown, Gel Green staining on the left and Coomassie Brilliant Blue Brilliant Blue staining on the right. Gel lanes not containing any samples visible with either of the stains or irrelevant for this figure have been cut from the images accordingly. Size in bp of relevant bands of the DNA-ladders (L1 and L2) are indicated. The position of DNA bands of competitor probes (C) and specific DNA probes (P) consisting of ArsR-BS are indicated. DNA-protein band positions are indicated (→) and red fluorescence of ACH is visible in gel areas coinciding with protein presence revealed by Coomassie Brilliant Blue Brilliant Blue staining (X). All samples and the running buffer contained 667.4 µM As(III). Samples contained 15 fmol of each DNA probe, 100 fmol of ACH (approx.. 4 µg) or lysate with a total protein content of 32 µg.

At the time of these experiments, cloning work was still under way to provide a His-tagged ArsR protein as a positive control. Cloning proved to be troublesome and while work was continued, alternative methods were considered. This led to the development of the Sepharose assay.

3.3.4 Sepharose assays for the visualisation of protein-DNA binding

As an alternative assay for the visualisation of ACH-DNA binding/unbinding in response to As(III), independent of gel electrophoresis, an ion-affinity based system was devised. Diethyl-aminoethyl–Sepharose (DEAE-Sepharose) is used widely in anion exchange protein purification techniques. Positively charged DEAE groups on

crosslinked, beaded agarose (Sepharose) bind negatively charged groups of proteins and DNA. Addition of salts changes the anion concentrations, competing with the protein groups or DNA as binding partners for the Sepharose. As previously shown by EMSA, the native form of ACH has a limited negative surface charge, leading to very slow migration through low percentage acrylamide gels. In denaturing electrophoresis techniques, such as SDS-PAGE, similar acrylamide percentages lead to significantly longer distances of travel in the same time (compare figure 3.3-15 and figure 3.3-6). Put briefly, native ACH was expected to have drastically lower surface charges available for binding of DEAE-Sepharose than ArsR-BS probes. Consequently, at increasing salt levels (NaCl), ACH was expected to be released from DEAE-Sepharose prior to ArsR-BS probes.

ACH and ArsR-BS probes were mixed in EMSA buffer and incubated in 1.5 mL tubes as described previously for EMSAs. Small quantities of DEAE-Sepharose CL-6B (Sigma-Aldrich) were added after incubation and tubes were inverted multiple times during a further 30 minute incubation step to allow for binding of protein / DNA to the Sepharose. The ideal concentration of protein / DNA identified by dilution series was found to be approximately 300 fmol DNA probe and 200 fmol of ACH in 1 mL 1x EMSA buffer with a reduced pH of 7.1. These ratios led to the formation of a relatively clear red fluorescent Sepharose pellet in the tubes upon 60 second centrifugation at 4500 x g. Addition of 250 mM NaCl led to the dissipation of ACH from the Sepharose after inverting the tube carefully and repeated centrifugation. Those samples containing ACH and ArsR-BS as a DNA probe showed less reduction in fluorescence of the pellet under the same conditions (figure 3.3-17).

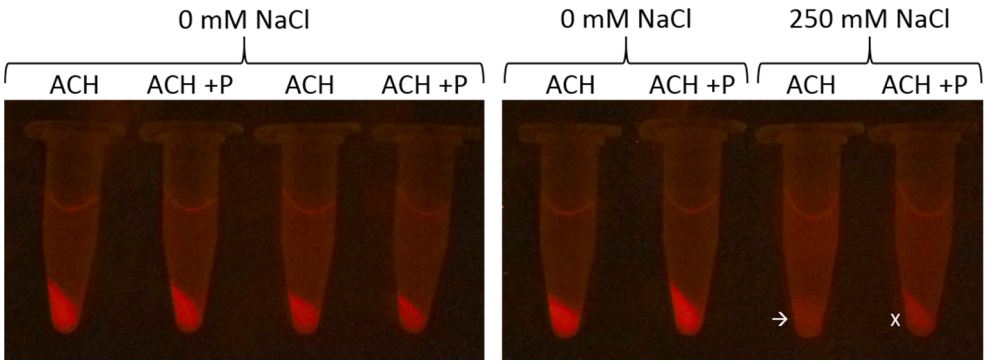

Figure 3.3-18: Sepharose assay of ACH and arsR-BS under varying NaCl concentrations.
300 fmol ArsR-BS as DNA probe (P) and 200 fmol of ACH were mixed with small amounts of Sepharose and incubated in 1 mL 1x EMSA buffer. Samples were centrifuged for 60 seconds at 4500 x g and an image was taken under blue light (left). 250 mM NaCl was added to half of the samples and all samples were inverted several times, then spun down as previously described. An image taken under blue light with the same camera settings is shown on the right. Reduction of fluorescence in the Sepharose pellet of the sample containing ACH alone is indicated by an arrow (→), less severe reduction of fluorescence in the sample containing ACH and DNA probe is indicated by the letter 'x'.

To make the somewhat subjective changes in fluorescence easier to compare, they were quantified with ImageJ software. Pictures taken from three experiments under identical camera settings and complete exclusion of perceptible ambient light were compared and numerical intensities calculated. Results showed that there was a slight decrease in fluorescence after the addition of EMSA buffer alone. Addition of NaCl in EMSA buffer to a final concentration of 250 mM lead to a decrease of pellet fluorescence to approx. 60% of the initial fluorescence for samples containing ACH and the ArsR-BS probe and to 10% for samples containing only ACH. Repetition of the experiments revealed standard deviations of 10 to 20 % above and below these values (figure 3.3-19, samples without As(III)). The experiments were repeated with As(III) added to final concentrations of 0.67 and 66.74 µM. Initial screening of the acquired images indicated no further decrease of Sepharose pellet fluorescence with increasing As(III) concentrations (figure 3.3-19 and figure 3.3-20). Concentrations above 500 µM As(III) led to diminishing fluorescence in all samples, regardless of the presence of any DNA probes. The experiments were performed with ACH + specific probe as well as with the competitor probe and no DNA at all. Samples with

ACH alone showed very little pellet fluorescence at any given As(III) concentration after the addition of 250 mM NaCl (figure 3.3-18).

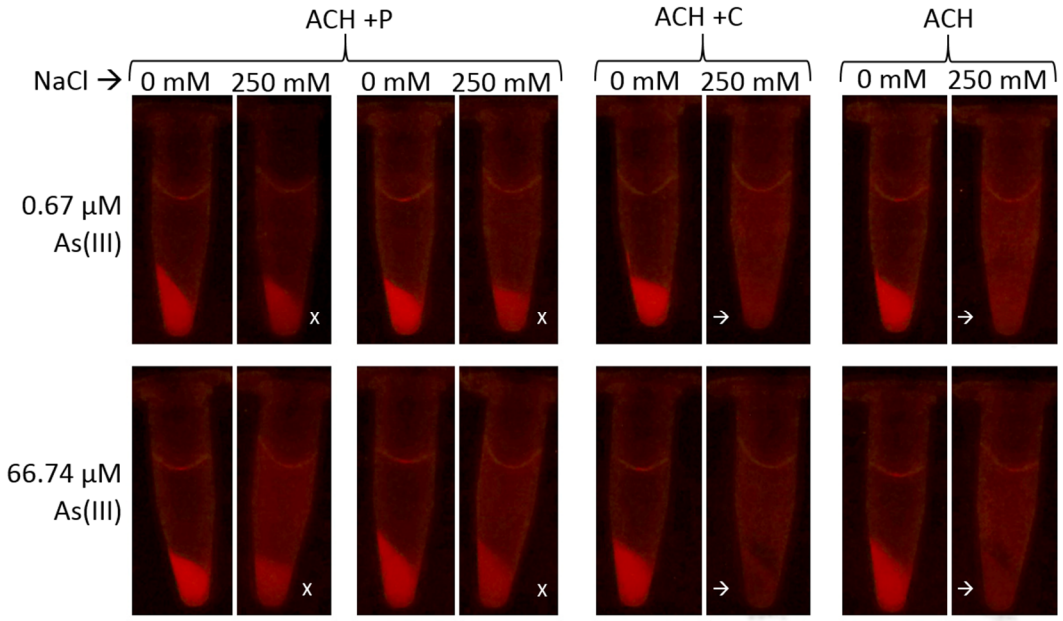

Figure 3.3-19: Sepharose assay of ACH, arsR-BS and competitor probe under varying As(III) and NaCl concentrations.

300 fmol ArsR-BS as DNA probe (P) and 200 fmol of ACH were mixed with small amounts of Sepharose and incubated in 1 mL 1x EMSA buffer. Samples were centrifuged for 60 seconds at 4500 x g and an image was taken under blue light (left). 250 mM NaCl and As(III) were added and samples were inverted several times, then spun down as previously described. An image taken under blue light with the same camera settings is shown on the right (250 mM columns) of each original image (0 mM columns). Reduction of fluorescence in the Sepharose pellet of samples containing ACH alone or ACH and competitor probe (C) is indicated by an arrow (→), less severe reduction of fluorescence in samples containing ACH and DNA probe (P) is indicated by the letter 'x'.

Images were analysed with ImageJ and results compared with the initial experiments without As(III). Pellet fluorescence in samples with ACH and ArsR-BS fluctuated around 60% of the initial fluorescence, matching the results without As(III) present regardless of actual As(III) content. The fluorescence in samples containing no specific DNA probe was so low that it fell entirely under the threshold set in the ImageJ software with the exception of ACH and 0.67 µM As(III) at 5% of the initial fluorescence (figure 3.3-19). It was concluded that the results were consistent with

the EMSAs in that ACH appeared to bind the specific but not the competitor DNA probes, leading to a stronger binding to the Sepharose if specific probes were present. However, the binding appeared to be non-responsive to As.

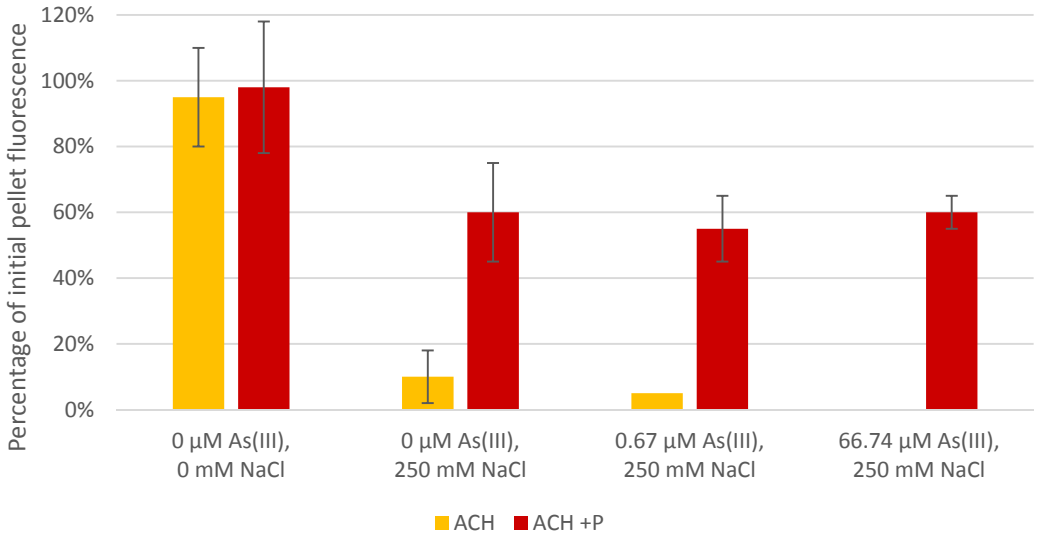

Figure 3.3-20: Fluorescence of Sepharose pellets binding ACH and / or the ArsR-BS DNA probe. Images of fluorescent Sepharose pellets, from experiments described before, were taken on a SafeView imaging system (Invitrogen), in absence of ambient light, with a Panasonic digital camera DMC-SZ7 at an aperture opening of f/3.1, ISO speed setting of 800 and an exposure time of 1 s. Cropped pictures were analysed with the ImageJ software. Average pixel intensity and size of the areas containing the fluorescent pellets was multiplied. Areas had been set by a threshold brightness and a threshold minimum size value to exclude noise. Fluorescence values of each tube were calculated before and after the addition of NaCl, As(III) or solvent alone. Fluorescence values are expressed as percentage of fluorescence prior to this addition. Where fluorescence was too low to form one connected area created by the software, separate spots of fluorescence in the area of the pellet were added together and multiplied with the average of all areas summarized. Error bars display one SD above and below the average of three independent replicates for samples without As(III) and duplicates for experiments with As(III) and ACH + probe (+P). No sufficiently bright pellet-fluorescence for calculation was detected for the sample containing ACH and 66.74 µM As(III).

3.3.5 Agarose as an alternative for polyacrylamide in EMSA

Since the results from the Sepharose assays were unsatisfactory and because it could not be excluded that the precast acrylamide gels had an influence on the results obtained for EMSAs, 3% agarose gels were prepared using the EMSA running buffer supplied with 667.4 µM As(III). Contrary to initial expectations, the resulting mobility shift showed crisp bands. Unfortunately, no release of the ArsR-BS probe

from ACH in response to As(III) was observed. The protein fractions – both, purified and crude lysates – did migrate further into the gel and shifting was more pronounced than on polyacrylamide gels. Purified ACH was visible as fluorescence and Coomassie Brilliant Blue stained bands between 766 and 916 bp as opposed to approximately 1500 bp on acrylamide gels. Shifted bands were located around 250 bp on agarose gels, less than half the 600 bp observed on acrylamide gels. Background from the lysates was spatially relatively separate from the shift bands, making the gel overall slightly easier to interpret. As previously reported for the PAGE mobility shift, Coomassie Brilliant Blue staining of the lanes containing crude lysates proved to provide little additional insight due to the high background (figure 3.3-20). The clear formation of shift-bands only in presence of the probe ArsR-BS indicates no As(III) responsiveness under agarose-EMSA-conditions at 667.4 µM. At this point it remained unclear whether this was due to impeded functionality of ACH or other factors. EMSAs in polyacrylamide gels with increased As(III) concentrations of up to 5 mM in the loading buffer did not lead to any perceptible changes in the shifting behaviour (not shown).

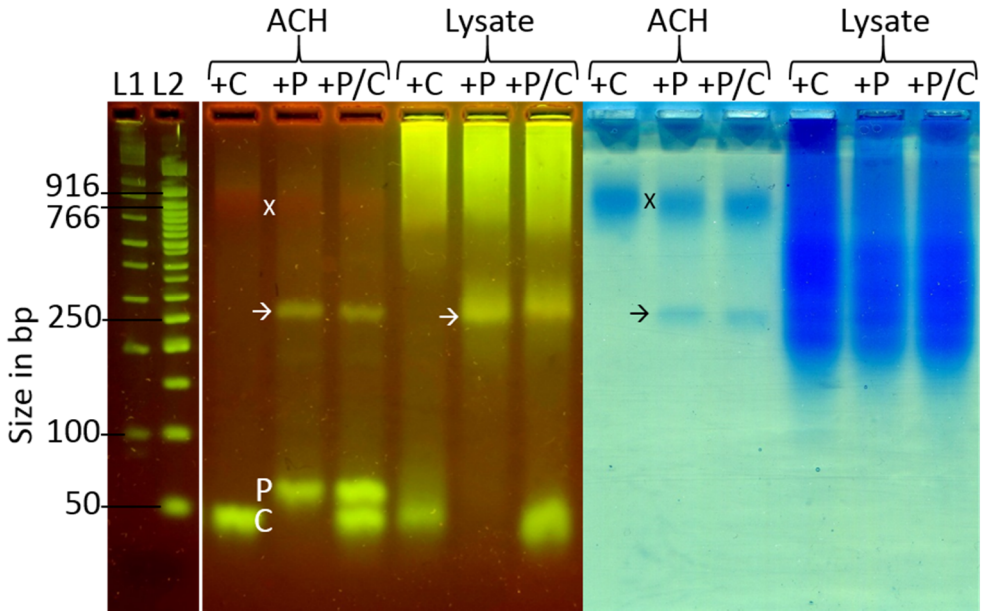

Figure 3.3-21: EMSA of ArsR-mCherry-His (ACH) and crude cell lysate on agarose gel.
Two stains of the same 3% agarose gel are shown, Gel Green staining on the left and Coomassie Brilliant Blue staining on the right. Gel lanes not containing any samples visible with either of the stains or irrelevant for this figure have been cut from the images accordingly. Size in bp of relevant bands of the DNA-ladders (L1 and L2) are indicated. The position of DNA bands of competitor probes (C) and specific DNA probes (P) consisting of ArsR-BS are labelled. DNA-protein band positions are indicated (→) and red fluorescence of ACH is visible in gel areas coinciding with protein presence revealed by Coomassie Brilliant Blue staining (X). All samples, the running buffer and the gel contained 66.74 µM As(III). Samples consisted of 15 fmol of each DNA probe, 100 fmol of ACH (approx. 4 µg) or lysate with a total protein content of 32 µg.

3.3.6 Second generation cloning and protein expression / purification

To determine whether the metal responsive properties of ArsR in ACH had been lost in the fusion protein, or if other factors were involved, a second generation fusion protein was created. In this new design, a linking sequence between ArsR and mCherry was to be introduced, securing separation of the two protein units and potentially restoring full ArsR functionality. A semi-flexible $(GGGGS)_3$ linker with optimised codon pairing, ensuring unimpeded expression levels, was chosen (Trinh et al. 2004). The linker sequence was introduced into the existing pET28-*arsR-mCherry-His* plasmid by MABEL (section 2.2.1.7). The resulting plasmid pET28-*arsR-(GGGGS)$_3$-mCherry-His* (figure 3.3-21) was transformed into *E. coli*

JM109(DE3), purified and validated by Sanger sequencing. DNA from the same preparation was used for the transformation of *E. coli* BL21(DE3). Protein expression was performed as described before (3.3.1), with similar results.

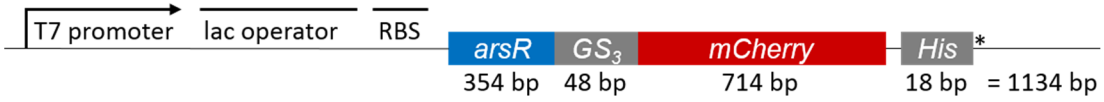

Figure 3.3-22: Schematic cloning scheme in 5'-3' orientation for fusion of *arsR* genes to *mCherry* in the pET28 vector introducing a $(GGGGS)_3$-linker.
Sequence elements originating from the vector are the T7 promoter, the *lac* operator, the RBS and the His tag as well as the stop codon (*). Relevant sequence lengths are given in bp below the corresponding parts and an approximate total of the sequence length for the fusion proteins is given on the right. The $(GGGGS)_3$-linker is labelled abbreviated to GS_3. The scheme is not to scale.

The fusion protein ArsR-$(GGGGS)_3$-mCherry-His (AGCH) showed intense fluorescence in liquid cultures and during all steps of purification as observed for ACH before. Purification via IMAC was performed exclusively without urea since the results reported for ACH had indicated no benefits from Urea addition. The expected protein mass was 42.2 kDa and SDS-PAGE of the various fractions from the IMAC purification indicated the presence of large quantities of a protein just under 46 kDa, making up almost the entire protein content of the purification fractions eluted at 0.1 M and 0.2 M imidazole (figure 3.3-22). These results are consistent with the purification results for ACH. Protein contents in the remaining two elution fractions were comparatively low. The SDS-PAGE showed relatively little contamination of the elution fractions with proteins of different size carried over after the washing step, as indicated by the absence of very strong bands apart from the ones described.

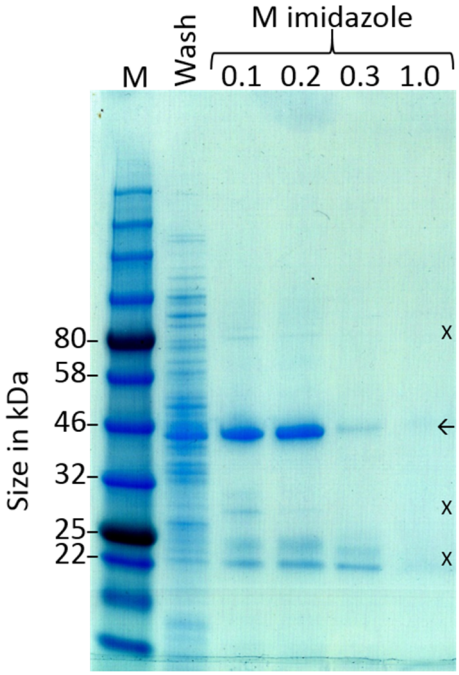

Figure 3.3-23: SDS-PAGE gels of AGCH purification and cultures.

Relevant sizes of the protein molecular weight marker (M) have been indicated. A: IMAC purification fractions for AGCH. Fractions shown from left to right are the wash fraction, elutions with 0.1 M, 0.2 M and 0.3 M imidazole as well as the final wash fraction with 1 M imidazole. A strong band is visible just below 46 kDa (←), indicating presence of AGCH at the expected 42.2 kDa. Some undesired protein contamination is carried over into the elution fractions (X). The equivalent of 10 µg total protein was loaded for the wash fraction and the 0.1 M elution fraction. The same volume as for the 0.1 M fraction was loaded of the remaining fractions to visualise differing protein concentrations.

3.3.7 ArsR-(GGGGS)$_3$-mCherry-His is arsenite responsive

Initial experiments were promising and showed the shifting previously observed for ACH. Neither P_{copA} probes, nor salmon sperm DNA led to a shift in AGCH travel, while incubation with the ArsR-BS probe led to a clear shift (figure 3.3-23).

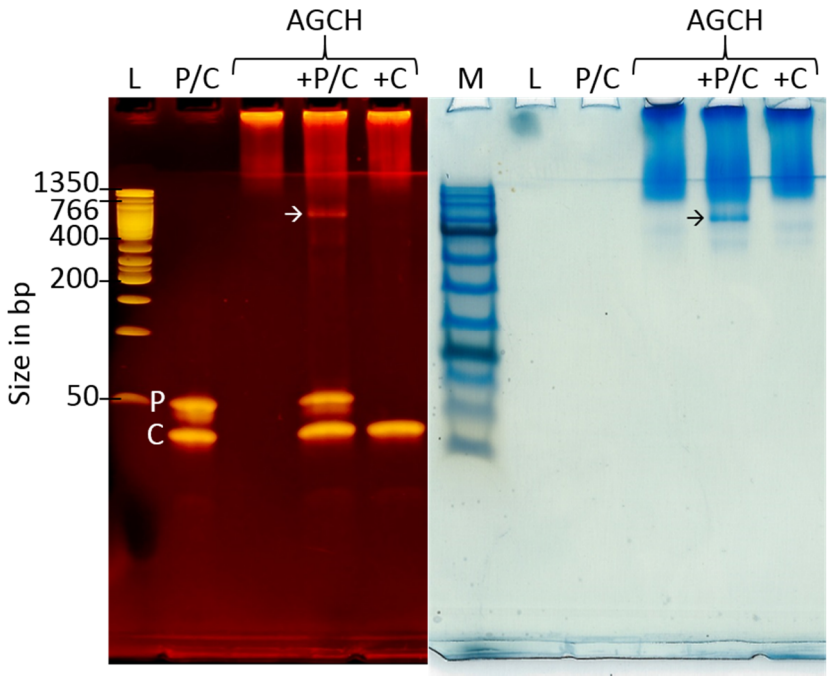

Figure 3.3-24: EMSA in acrylamide gel of AGCH with specific and competing DNA probes. Two stains of the same gel are shown, Gel Green staining on the left and Coomassie Brilliant Blue staining on the right. Gel lanes not containing any samples visible with either of the stains have been cut from the images accordingly. Size in bp of relevant bands of the DNA-ladder (L) are indicated. Note that the native protein samples did not enter the gel as they would in SDS-PAGE and that protein standard labels have been left out for that reason. The position of DNA bands of competitor probes (C) and specific DNA probes (P) consisting of ArsR-BS are labelled. DNA-protein band positions are indicated (→) and red fluorescence of AGCH is visible in gel areas coinciding with protein presence revealed by Coomassie Brilliant Blue staining. Approximately 10 fmol of each DNA probe and 100 fmol of AGCH were mixed and pre incubated in 1x EMSA-loading buffer before loading into the appropriate gel wells.

At this stage of the project the His-tagged, unfused ArsR (ArsR-His) was successfully cloned and could be incorporated into the following experiments. Since the capacity of AGCH to specifically bind the operator region of P_{ars} had been demonstrated, metal-induction testing commenced. Samples were incubated as described previously with various concentrations of As(III) in the form of sodium metaarsenite (referred to as As(III)) and phenylarsine oxide (PAO) as well as As(V) in the form of sodium arsenate. Shifts shown in figure 3.3-25 and figure 3.3-26 contained ArsR-BS-TC as a DNA probe. ArsR-BS-TC is a shortened version of the

ArsR-BS probe with an added poly(10T)poly(10C) tail at the 5' end (see materials and methods for a full sequence ,Sun et al. 2012),.

Gel shifts showed an intensity reduction of the protein-DNA band in response to As(III) and PAO but not As(V) (figure 3.3-25). The response to PAO was more pronounced and required drastically lower concentration when compared to As(III). Figure 3.3-25 shows an EMSA in a 5% polyacrylamide gel for ArsR-His. Shift bands were observed to run between 100 bp and 150 bp, lower than for AGCH-shifts (400 bp). ArsR-His was predicted to have a molecular mass of 14.1 kDa, less than a third of the 46.2 kDa AGCH fusion protein. As(III) concentrations of 50 µM and above were observed to reduce the intensity of the shift band, with the band becoming indiscernible at 50 - 500 µM for the Gel Green stain. The Coomassie Brilliant Blue stain revealed corresponding bands up to similar concentrations with their intensity decreasing drastically as As(III) concentrations increased. The addition of PAO led to the same effects, but at lower concentrations. 0.5 µM PAO led to a decreased band intensity in both stains and at higher concentrations (5 µM and 50 µM) the band became invisible. Overall, the PAO concentration required to disrupt the binding of ArsR-His with the ArsR-BS-TC probes was approx.. 10 times lower when compared to As(III).

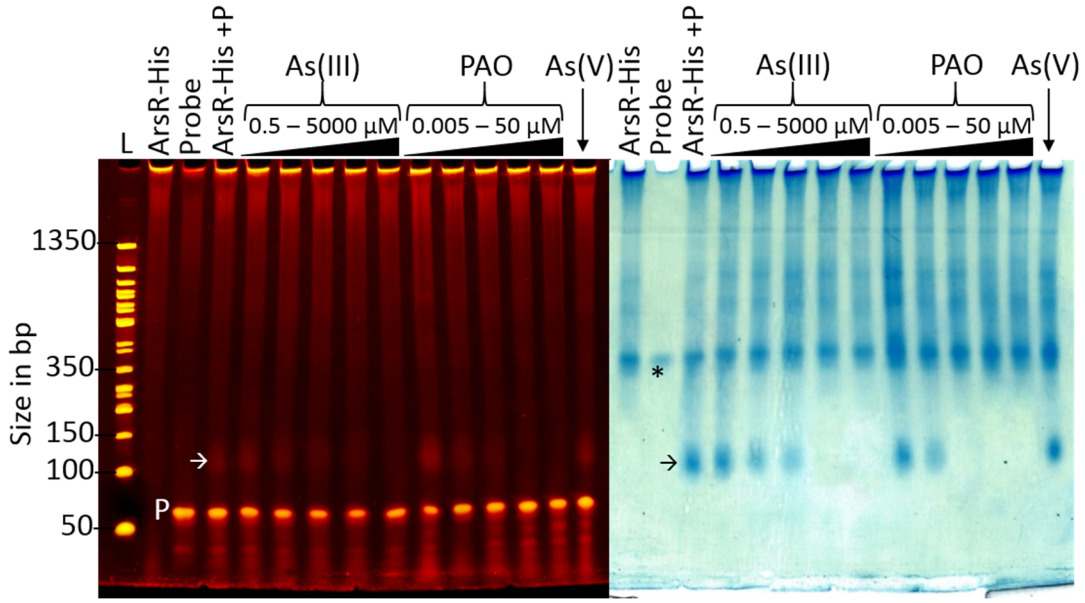

Figure 3.3-25: EMSA of ArsR-His proteins and ArsR-BS-TC DNA probes.
Two stains of the same 5% TBE-polyacrylamide gel are shown, Gel Green staining on the left and Coomassie Brilliant Blue staining on the right. Size in bp of relevant bands of the DNA-ladder (L) are indicated. The position of DNA bands of DNA probes consisting of ArsR-BS-TC are marked by the letter P. Bands consisting of ArsR-His bound to the probe are indicated (→) and coincide between the stains. As(III) as sodium metaarsenate was added to samples marked accordingly at concentrations of 0.5 µM, 5 µM, 50 µM, 500 µM and 5000 µM. Phenylarsine oxide (PAO) was added to the samples at concentrations of 0.005 µM, 0.05 µM, 0.5 µM, 5 µM and 50 µM. As(V) as sodium arsenate was added as a control at 5000 µM. All potential inducers were added to the binding mixtures prior to sample incubation. Reactions contained 10 fmol of DNA probe and 200 fmol of ArsR-His where applicable. All samples contained BSA and salmon sperm DNA at 50 µg/mL each. BSA bands visible in the Coomassie Brilliant Blue stain are marked by an asterisk.

A similar gel shift was performed with AGCH. Results were consistent with the observations based on ArsR-His but metal responsive release of the DNA probe required slightly increased metal concentrations when compared to shifts of ArsR-His, as shown in figure 3.3-25. While 500 µM As(III) had led to a complete disappearance of a shift-indicating band on the Coomassie Brilliant Blue stain of the ArsR-His shift, the AGCH EMSA still showed a clearly discernible band under the same conditions. Similarly, PAO at 0.5 µM concentration seemed to lead to total DNA unbinding of ArsR-His, but not AGCH.

As had been observed before for ArsR-His, AGCH released the ArsR-BS-TC probe upon PAO concentrations approx. 10 times lower than As(III). When comparing the results obtained with ArsR-His and AGCH, it became apparent that the concentrations of PAO or As(III) required to disrupt ArsR-His binding were approx. 10 times lower than those required to disrupt AGCH binding.

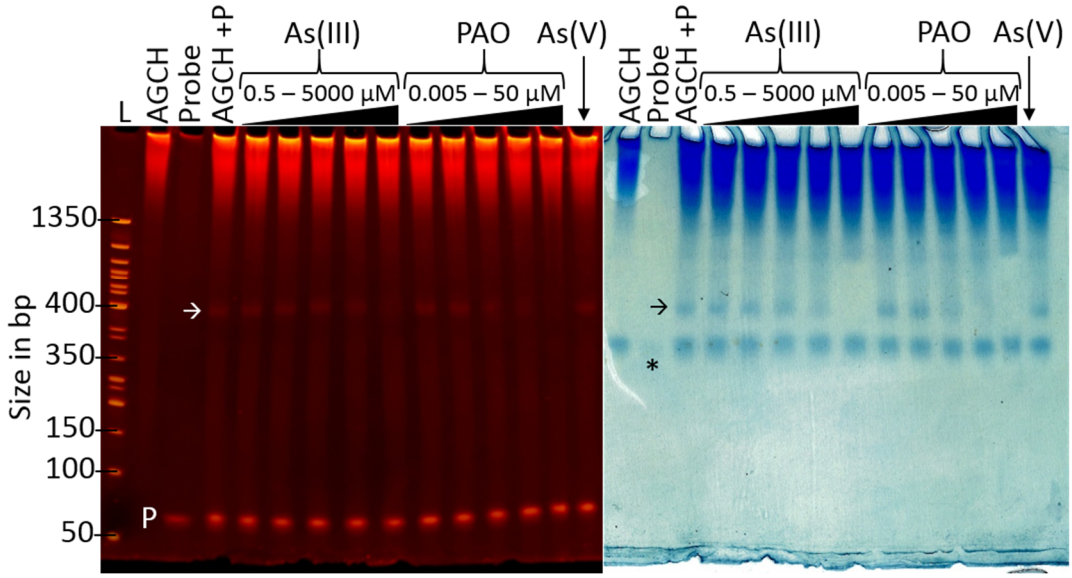

Figure 3.3-26: EMSA of AGCH proteins and ArsR-BS-TC DNA probes.

Two stains of the same 5% TBE-polyacrylamide gel are shown, Gel Green staining on the left and Coomassie Brilliant Blue staining on the right. Size in bp of relevant bands of the DNA-ladder (L) are indicated. The position of bands of DNA probes consisting of ArsR-BS-TC are marked by the letter P. Bands consisting of AGCH bound to the probe are indicated (→) and coincide between the stains. As(III) as sodium metaarsenate was added to samples marked accordingly at concentrations of 0.5 µM, 5 µM, 50 µM, 500 µM and 5000 µM. Phenylarsine oxide (PAO) was added to the samples at concentrations of 0.005 µM, 0.05 µM, 0.5 µM, 5 µM and 50 µM. As(V) as sodium arsenate was added as a control at 5000 µM. All potential inducers were added to the binding mixtures prior to sample incubation. Reactions contained 10 fmol of DNA probe and 200 fmol of AGCH where applicable. All samples contained BSA and salmon sperm DNA at 50 µg/mL each. BSA bands visible in the Coomassie Brilliant Blue stain are marked by an asterisk.

3.3.8 Only non-porous carrier materials enable AGCH release

In order to find a suitable carrier material for the AsGard test strip, various materials were tested. The most basic requirement was that the AGCH protein should be easily washed off the surface or be mobilised by capillary forces within the material. Initially three types of paper were tested: ordinary office paper (recycled, white, 80 g/m^2, 5226498, OfficeDepot), filter paper (1001-185, Whatman) and blotting paper (GB003, Whatman). 5 µg of AGCH in 3 µL EMSA buffer were added to the paper. Before the protein solution dried, the paper was immersed into 1x TE buffer and washed under light agitation for 15 minutes at RT. The wet paper was then transferred onto the SafeView imaging system and pictures of fluorescence from AGCH were taken. All paper types showed clear fluorescence in the areas where AGCH had been applied, without any perceptible difference in brightness compared to controls that had been imaged directly after application of the protein. Extended washing, up to the disintegration of the carrier material, did not lead to any reduction in fluorescence. Consequently papers were treated with blocking agents used in blotting applications, namely BLOTTO (ThermoFisher) and Denhardt's solution. Still, fluorescence remained unchanged before and after washing. In a new approach blotting membranes were tested as a possible support. Hybond N+ (nylon), PVDF and Nitrocellulose membranes (all Amersham) were blocked with Denhardt's solution for 2.5 h and AGCH was applied as previously described. PVDF membranes showed no decrease of fluorescence after 30 minutes or 3 days of continuous washing under agitation. Hybond N+ membranes showed some reduction of fluorescence after three days, but none after 30 minutes. The only reduction in fluorescence that was somewhat clear was achieved with the nitrocellulose membrane. 30 minute washing led to a perceptible decrease and after 3 days fluorescence had vanished almost completely. Figure 3.3-27 shows an exemplary set of pictures for each membrane type employed. Results were deemed unsatisfactory and the focus shifted onto less porous materials.

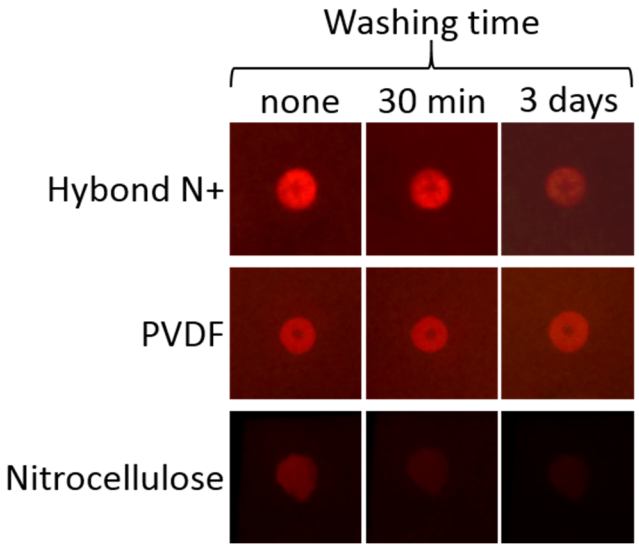

Figure 3.3-27: Blotting membranes showing fluorescent AGCH dots before and after washing. Each dot contains 5 µg of AGCH in EMSA buffer. Pictures were taken with identical camera settings after application, 30 minute wash and 3 day wash. Pictures shown are exemplary for multiple experiments carried out. Washing was performed at RT under agitation in TE buffer.

3.3.9 PMMA as a carrier material

Poly(methyl methacrylate) (PMMA) is widely known as acrylic glass or Perspex and researchers have shown the direct immobilisation of ssDNA probes on non-modified plastic surfaces – amongst which was PMMA – with UV irradiation and a simple poly(T)poly(C)-tag (TC_{tag}) (Sun et al. 2012). The ArsR-BS sequence was altered to incorporate the TC_{tag}, consisiting of 10 T and 10 C bases, at the 5' end of the coding strand, resulting in ArsR-BS-TC (figure 3.3-27). Refer to section 2.2.4.5 in material and methods for a full sequence including the TC_{tag}.

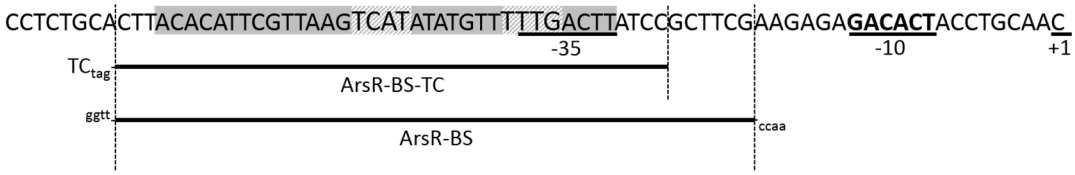

Figure 3.3-28: Nucleotide sequences of the ArsR binding DNA probes.
ArsR binding sites after Xu et al. have a hatched background. The shaded sequence including hatched areas indicates the ArsR binding region after Xu et al. as defined by DNase I foot printing. The -35 and -10 promoter regions as well as the transcription start (+1) are underlined. The coverage of the given sequence by both iterations of DNA probes is shown by black bars and indicated accordingly. Surface binding tags (TC_{tag}) (see 2.2.4.5 for full sequence) and single strand overhangs are shown at their respective positions (Xu et al. 1996).

PMMA (PLEXIGLAS XT, Evonik Industries AG, Germany) was cut into 20x70x2 mm slides. Double stranded and ssDNA were applied and bound according to the method published by Sun et al., except that the volumes were increased drastically to 0.5 µL to allow for manual liquid handling (Sun et al. 2012). To visualise binding of the DNA, slides were stained for 20 minutes in a 6x Gel Green staining bath after washing and imaged on the SafeView imaging system with a Sony DSC-RX100 Mk1 digital camera (figure 3.3-28). Slides were then returned to the wash bath and left under agitation overnight. Images taken on the next day showed no perceptible decrease of fluorescence. Different shapes of DNA application were tested, amongst which was a stripe shape, produced by slowly pulling a pipette tip containing Ars-BS-TC in spotting buffer transversally over the slide. However, none of the alternative shapes gave rise to fluorescence stronger than the dot-shape, application was slower and precise volume control impossible. Thus, experiments regarding DNA application shape were stopped.

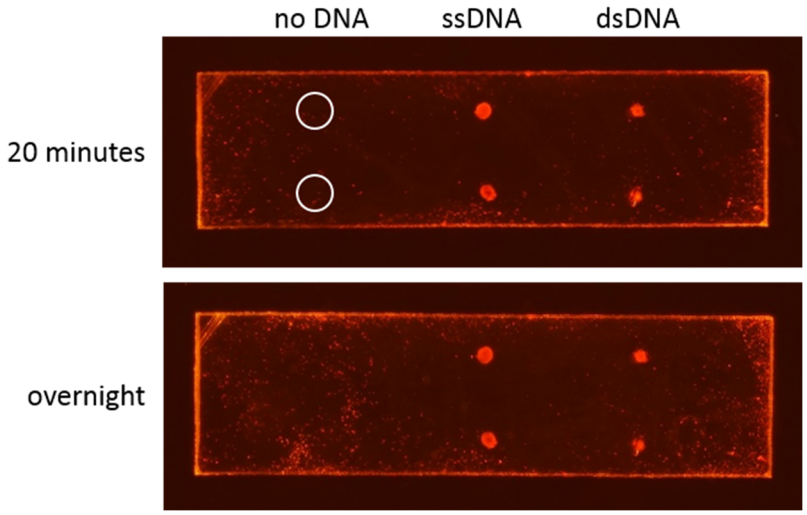

Figure 3.3-29: Fluorescence of DNA bound to PMMA slide and stained with Gel Green.
0.5 µL spotting buffer containing 25 µM ArsR-BS-TC DNA was applied to the slide, allowed to air dry and bound via UV irradiation. White circles mark the area where buffer containing no DNA was added. The same slide was imaged after 20 minutes and overnight staining/washing under agitation. Replicates are shown above each other on each slide. Exposure time was 6 s at an aperture setting of f/1.8 and ISO 125.

To determine whether protein binding to immobilised DNA was possible, new slides were prepared and DNA bound before AGCH protein was applied. The protein was applied in EMSA buffer and alternatively in protein binding buffer (PBB, 2.2.6.1). Both slides were imaged immediately and one slide was allowed to air dry at RT for approximately 45 minutes before washing with KP-T buffer (10 mM potassium phosphate pH 6.0, 0.05% (w/v) Tween 80). The second slide was washed immediately. Both slides were imaged after washing and neither showed any remaining fluorescence for any of the application points. Gel green staining revealed that DNA was still present. Subsequently, an incubation period of 20 minutes after the addition of AGCH to the slides was introduced. During this incubation time it was ensured that the protein spots did not dry out by ensuring that slides were not exposed to temperatures above 20°C, close by light sources or draughts. Since the application of 0.5 µL dots required the use of a standard handheld microliter pipette (Research plus, Eppendorf), 1 µL was tested as an alternative. 1 µL could be applied with the help of a stepper pipette (Multipette M4, Eppendorf), greatly increasing dotting speed and volume accuracy without touching the slide surface. A slide holder

was fabricated as a positioning device and dots were applied following a system of coordinates, as DNA-dots were barely visible once the DNA was bound and the slides had been washed and dried. Images were taken after protein dotting, incubation and washing and are shown in figure 3.3-29. Fluorescence somewhat decreased during the 20 minute incubation and began to form a "donut" shape with reduced intensity in the centre. No fluorescence was detected after washing.

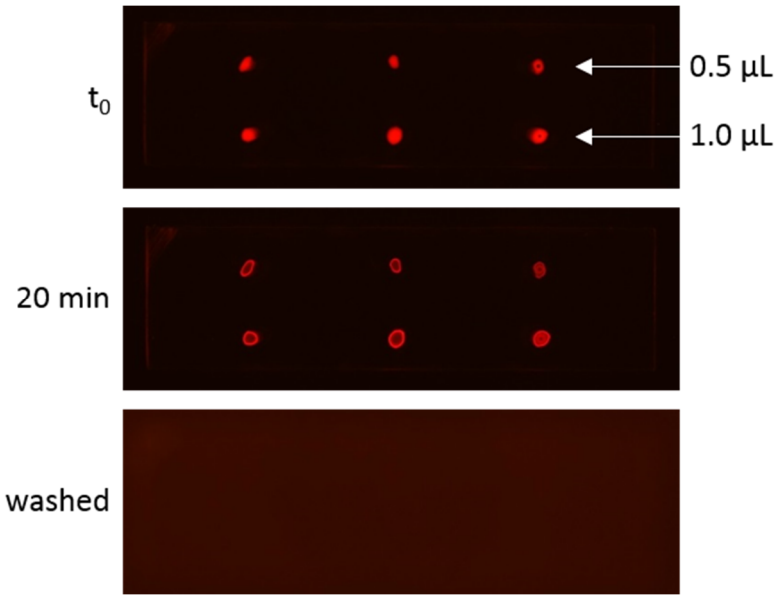

Figure 3.3-30: Fluorescence of AGCH applied to a PMMA slide with ArsR-BS-TC.
AGCH in EMSA buffer (1 µg/µL) was spotted onto a PMMA slide with immobilised ArsR-BS-TC probes. The top row was spotted with 0.5 µL each, the bottom row with 1 µL. Images of protein fluorescence were taken directly after application of AGCH and after 20 minute incubation without drying as well as after a 20 minute wash with KP-T buffer under careful agitation. Note that the last image has been taken with a longer exposure time to ensure no fluorescence was missed due to under-exposure. Exposure times for the first two pictures were 2 s, the bottom picture was taken with 10 s exposure time. Aperture was f/1.8 and ISO 125.

The experiment described above was repeated and slides were washed for 10 minutes in KPB buffer (50 mM potassium phosphate pH 6.7, 50 µg/mL salmon sperm DNA, 40 mM NaCl) under careful agitation instead of KP-T buffer. This led to some weak fluorescence remaining on the slides, localised at the immobilised DNA. As observed in the previous experiment, the fluorescence began to appear in a donut shape during incubation and remained in this shape after washing. Exposure time had to be

increased to 10 s at maximum aperture opening (f/1.8) to image the remaining fluorescence, increasing the background from light passing the filter.

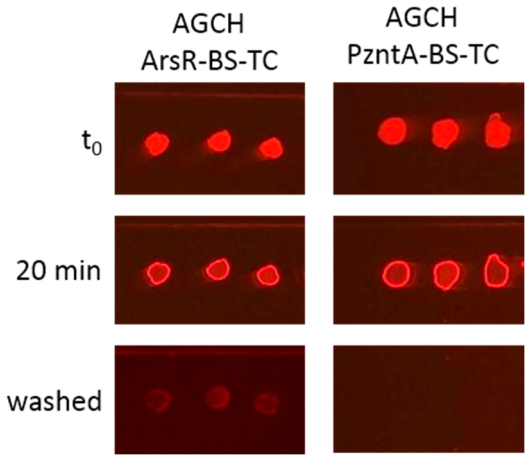

Figure 3.3-31: AGCH fluorescence on PMMA slides endures washing only in presence of ArsR-BS-TC probes.

AGCH proteins were applied to a slide containing immobilized ArsR-BS-TC probes and P_{zntA}-BS-TC probes as control. The slides were imaged after application, a 20 min incubation period and a 10 min wash in KPB buffer. Note that different exposure times are shown to show the localisation of the fluorescence more clearly. Exposure times were 5 s for the first two rows of images and 10 s for the last images. Aperture was f/1.8 and ISO 125.

Several experiments were undertaken to increase the level of fluorescence remaining after washing. Protein was buffered in PBB instead of EMSA buffer, incubation times were extended and slides were rinsed with 70% ethanol and ultrapure water prior to DNA immobilisation. None of these alterations increased the fluorescence after the washing step. In fact, changing the protein buffer to PBB decreased the fluorescence slightly, although differences were too small to draw any conclusions from this. Due to time constraints, it was decided to commence with testing the metal responsive properties of the AGCH immobilisation.

3.3.10 AsGard – an arsenite responsive PMMA-slide assay

Following the optimisations previously described, slides were prepared as follows:

1. Spots of 1 µL spotting buffer containing 25 µM ArsR-BS-TC were applied onto a fresh PMMA slide without touching the surface with the pipette tip. Spots were allowed to dry completely.
2. Slides were exposed to UV radiation at 254 nm with a power of 3 mW/cm^2 for 10 minutes.
3. Slides were washed under agitation in 0.1x standard saline citrate (SCC) with 0.1 (w/v) SDS for 10 minutes and thoroughly rinsed in deionised water. Slides were dried by air flow.
4. AGCH was applied in 1 µL EMSA buffer with BSA and salmon sperm DNA, slides were incubated for 20 minutes without allowing the spots to dry.
5. Excess protein was washed off in 10 mL KPB under careful agitation for 5 minutes (pre-wash).
6. Slides were transferred into 10 mL KPB containing inducers at appropriate concentrations and kept under careful agitation, interrupted for regular imaging (metal wash).

This technique gave rise to PMMA slides showing a time and As(III) but not As(V) concentration dependent decrease in AGCH fluorescence. Differences between those slides that had been washed in higher As(III) concentrations and those with zero or very low concentrations were most pronounced for very short washing times of 5 + 5 minutes (pre-wash and metal wash). Washing for longer than 25 minutes (including the pre-wash) led to a loss of fluorescence on all slides, independent from metal concentrations in the buffer. Choosing a pre-wash duration shorter than 5 minutes sometimes led to excess AGCH remaining on the slide surface, visible as fluorescent “clouds” (figure 3.3-31).

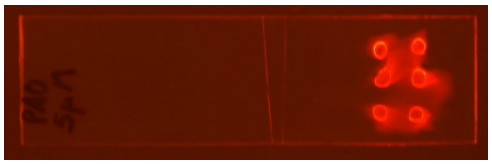

Figure 3.3-32: Insufficient pre-washing leads to the formation of fluorescent “clouds”.
The slide was washed for 2 minutes in KPB under careful agitation. Image was taken with an aperture of f/1.8, ISO 125 and an exposure time of 10 s.

As(III) concentrations of 0.067 µM (approx. 5 ppb) did not lead to a fluorescence decrease that was notably faster than in those experiments without As(III). However, increasing the As(III) concentration to 0.134 µM (approx. 10 ppb) led to an increased speed of fluorescence depletion and concentrations of 0.267 µM (approx. 20 ppb) and above led to no fluorescence being visible after 5 minutes of metal washing. As a control, slides were washed in buffers containing 1.335 µM As(V) and no perceptible difference to the behaviour of slides that had been kept in metal free KPB was seen (figure 3.3-32).

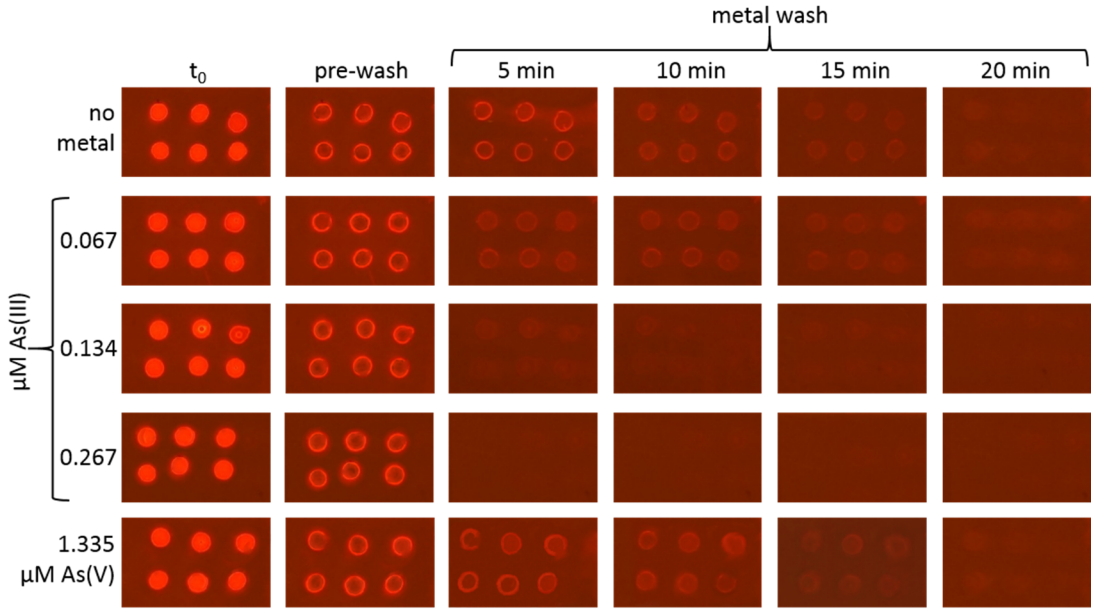

Figure 3.3-33: As(III) responsive decrease of visible AGCH fluorescence on PMMA slides.
For each condition ArsR-BS-TC was immobilised on a fresh PMMA slide and AGCH protein applied as previously described. All slides were incubated for 20 min and unbound proteins were washed off in KPB for 5 min. Subsequently, slides were transferred into KPB containing As(III) at 0.067 µM, 0.134 µM or 0.267 µM, As(V) at 1.335 µM or no added metal. Images were taken directly after protein addition (t_0), after the initial pre-wash and every 5 minutes thereafter. All images were taken with an aperture of f/1.8, ISO 125 and an exposure time of 10 s. All wash-steps were performed under careful agitation on a rotary table.

The images obtained were analysed with ImageJ software in an attempt to quantify the results. The analysis provided some numeric data, showing the relatively linear decrease of fluorescence after the pre-washing. Additionally, it revealed a slightly faster reduction in the 0.067 µM As(III) samples. This difference had been too weak to be seen clearly in the photographs (figure 3.3-33). For higher concentrations or longer washing times, the analysis did not provide any additional insights. This was partially due to the high background light levels in consequence of the long exposure times.

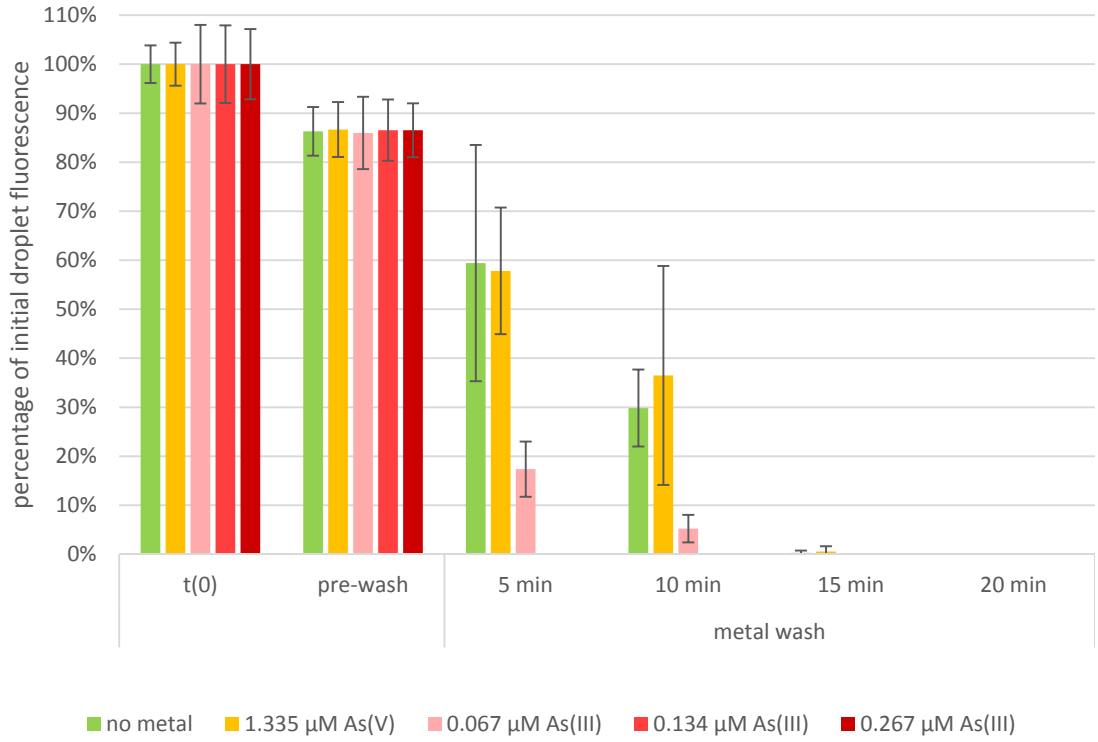

Figure 3.3-34: ImageJ analysis of the fluorescence images shown in figure 3.3.10-II.
Average pixel intensity and size of the fluorescent areas were multiplied. Areas had been set by a threshold brightness and selected manually to exclude noise. Bars show the fluorescence intensity as a percentage of initial fluorescence after protein application. Error bars display one SD above and below the average of 6 dots per sample. Absence of bars indicated the brightness of fluorescent dots had fallen below the threshold value necessary to exclude background light.

To reduce exposure times, the fluorescence signal had to be increased. One way of achieving this was thought to be an increase in binding area and structure. Consequently, new PMMA slides were sanded with fine grit sanding paper (Wetordry Tri-M-ite P600, 3M, Bracknell) to an opaque finish. Slides were cleaned thoroughly and rinsed with deionised water before continuing with the preparation process as previously detailed. Results were unsatisfactory in that fluorescence was washed off the slide more easily, with almost no fluorescence being left after 10 minutes of washing in KPB. Additionally, released AGCH had a tendency to adhere to the PMMA surface and the previously mentioned “cloud” formation was intensified when compared to smooth PMMA (figure 3.3-34). Due to this, the slightly increased initial fluorescence did not present any advantage (note that images were taken with half the exposure time of the previously presented experiment).

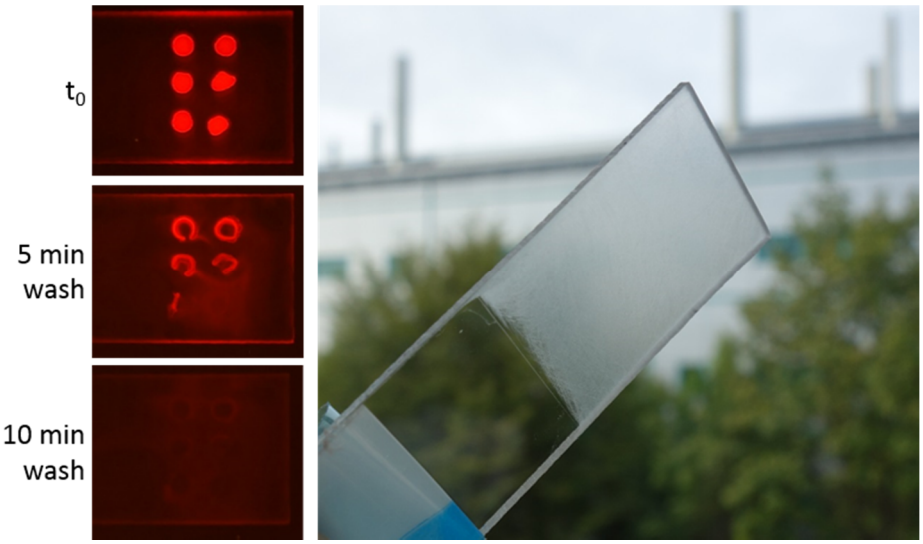

Figure 3.3-35: Fluorescence on sanded PMMA slide.
Fluorescence was imaged directly after AGCH application (t_0) and after 5 and 10 minutes of washing in KPB. The right hand side of the figure shows the appearance of a sanded and cleaned slide. Fluorescence images were taken at an aperture setting of f/1.8, ISO 125 and with 5 seconds exposure.

3.3.11 Approaches for long term storage of AsGard

When the protein spots of AGCH on the PMMA slides were allowed to dry completely before or after washing or at any other point of the experiments, fluorescence initially only decreased slightly but AGCH no longer stayed bound to the DNA probes and was washed off instantly upon rehydration. Sugars have been widely used for the preservation of dried proteins. Trehalose and lactose were chosen as candidate drying protectants. AGCH in EMSA buffer was supplemented with trehalose to final concentrations of 30, 20 and 10% (w/v) and lactose to 10% (w/v) (Yamamoto & Sano 1992; Lippert & Galinski 1992). Some publications have used Tween to further stabilise proteins against dehydrational and thermic stress (Costa-Silva et al. 2011). Consequently, as a further means to enhance the AsGard assay, some samples were supplied with Tween 80 to a final concentration of 0.1% (v/v).
To elucidate the influence of sugars and Tween on the DNA-AGCH interaction on the slide, the solutions were spotted onto PMMA slides with immobilised DNA probes as previously described. Slides were imaged after spotting, incubation and after being washed for 10 minutes in KPB. Spots that contained lactose or trehalose

showed decreased attachment of AGCH to the DNA probe. Highest attachment after washing was found to be present in samples without sugar or Tween but containing BSA and salmon sperm DNA (figure 3.3-35). This effect had been observed previously and was the reason for inclusion of BSA and salmon sperm DNA in the AsGard assays previously described.

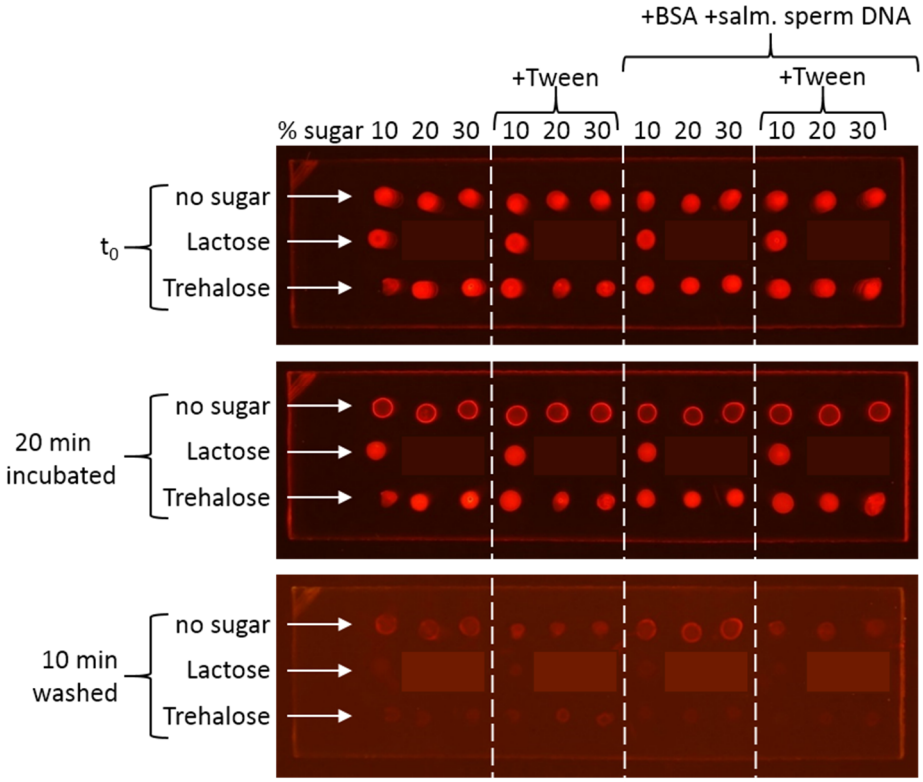

Figure 3.3-36: Influence of lactose, trehaolse, Tween, BSA and salmon sperm DNA on AGCH binding to ArsR-BS-TC on PMMA slides.

Shown are images of the same slide directly after sample application (t_0), after 20 min incubation and 10 min washing in KPB. Dots containing no sugar, lactose or trehalose are shown above each other on each slide as marked on the left. Samples without sugar were supplied with an equivalent amount of ddH_2O. Lactose was only added to a final concentration of 10% as solubility did not allow higher concentrations. Each concentration was tested without further additions or combined with 0.1% Tween 80, BSA and salmon sperm DNA or both (accordingly indicated above the slide pictures). Note that the unused lactose concentration positions contained more water samples and have been covered in this figure to make viewing easier. The exposure time for the washed slide was increased from 5 to 10 seconds.

3.4 Discussion

3.4.1 Cloning, protein expression and purification

After initial difficulties cloning the various gene-reporter combinations, the test and expression cultures showed good growth after induction, indicating that the protein fusions confer a relatively low fitness cost to *E. coli*. In case of ACH and AGCH the *de-novo* synthesis of the desired protein fusions could be confirmed by the clearly visible red colour developing in the cultures after induction or sufficient growth time in auto-induction medium. The SDS-PAGE gels showed strong expression for all protein fusions and allowed the setting of an ideal harvesting point at 3 h after induction at an OD_{600} around 2.1 for the IPTG inductions. Later auto-inductions required longer growth times but overall protein yields were increased.

The initial cloning scheme and the fusion proteins it gave rise to had been conceived considering a range of aspects. ArsR functions as a dimer and as such it was considered a potential source for complications to use strongly multimeric proteins as a marker. mCherry is a monomer and EYFP a weak dimer, decreasing the chance of unfavourable aggregation of fusion proteins. EYFP was expected to be of superior brightness under blue light (470 nm) since its excitation profile spans from approx. 450 nm to 540 nm peaking around 500 nm with a pronounced shoulder at 470 nm. In contrast, mCherry has its excitation maximum at 587 nm and excitation below 500 nm is normally reported as minimal (Day & Davidson 2009). The very weak or totally absent fluorescence in samples containing protein fusions with EYFP might have various causes. EYFP might be more susceptible to misfolding or denaturation than mCherry when fused to ArsR and a His-tag. While the light wavelength of the SafeView imager had been considered ideal for EYFP use and was in fact a reason for the inclusion of EYFP in the project, the bandpass of the amber filter unit of this imaging system is unknown. It is possible, that for EYFP the emission is affected adversely by the filter. Since the mCherry clones were superior in fluorescence visibility, protein yields, apparent stability and daylight visibility, EYFP experiments were not prioritised and later discontinued. mCherry however might not be the ideal marker protein despite its high stability. A darker dye giving a higher contrast would

be beneficial for visibility in daylight. AmilCP is a recently described chromoprotein from the coral *Acropora millepora* and has been considered as a candidate for a reporter protein (Alieva et al. 2008). Its expression results in an intense dark blue / purple colour. The sequence is of entirely natural origin and all sequence modifications such as removal of restriction sites and codon optimization for *E. coli* have been performed by the Uppsala University iGEM team of 2011 (Alieva *et al.* 2008, personal communication Erik Gullberg). While the chromo-properties are highly desirable, the tetrameric nature of amilCP might present a problem. Since the results obtained with the mCherry fusions as discussed later were not as envisaged, no amilCP fusions were prepared. Should the binding efficiency of the PMMA assay be increased in the future, these fusions might become desirable again. The initial strategy to combine mCherry and EYFP with ArsR incorporated no linker and instead was based on the assumption that the flexible C-terminal tail of the ArsR protein would act as a linker (1.2.3.1). Based on the EMSA data, a flexible (GGGGS)3-linker was later introduced between ArsR and mCherry to reduce potential steric hindrance in ArsR dimer formation and function.

3.4.2 Non-radioactive mobility shifts

A crucial tool for visualisation of protein-DNA interactions has been and still is the electromobility shift assay. While other methods, such as surface plasmon resonance, have been developed and can have superior sensitivity. The EMSA remains a relatively easy to perform, low cost and sufficiently sensitive method (Hellman & Fried 2007). However, the EMSA is not without limitations. High salt concentrations in buffer, gel and sample increase the conductivity and lead to rising temperatures, potentially disturbing any DNA-protein complexes. Use of different buffers for gel running and sample incubation can lead to unfavourable influences on complexes (Fried 1989). Extended electrophoresis times can have detrimental effects and lead to dissociation of complexes. This is due to samples not being at chemical equilibrium during the electrophoresis, sometimes leading to dissociation of complexes while many proteins have been described to bind significantly more stable to their DNA partners under EMSA conditions (Fried & Liu 1994; Fried & Bromberg 1997; Vossen & Fried 1997). With increasing regulations and the latent pressure to reduce

the use of radioactive techniques to an absolute minimum, the traditional radiolabelling of DNA in EMSAs has become one of their major disadvantages. Non-radioactive detection of the nucleic acids with post-electrophoretic staining has been described in the past (Jing et al. 2003). The EMSA results in this work show clearly the sensitive imaging of protein-DNA complexes via post-electrophoretic staining with safe DNA stains and additionally staining of the protein fractions on the gels via modified Coomassie Brilliant Blue staining. This combined approach removes the usual uncertainty about the composition of bands visible through nucleic acid imaging techniques such as radiolabelling. Most traditional EMSA techniques rely on the imaging of the DNA fractions alone, leaving some hypothetical doubt about the localisation of proteins and perhaps more importantly the ratio of proteins involved in the observed complexes. The non-radioactive EMSA established as part of this work and inspired by a publication by Jing *et al.* shows, beyond reasonable doubt, that shifted bands consist of DNA and protein and that these complexes form a relatively small fraction of the total protein amount invested.

In the specific case of ACH and AGCH it remains largely unclear if these uninvolved proteins are unable to bind the DNA probe due to the conditions in the binding buffer, if they release them after gel loading or if large portions of the proteins are otherwise incapacitated as a consequence of sonication, purification and storage. The inclusion of crude extracts in the gel shifts of ACH indicates that the purification steps and storage conditions / duration have little or no influence, yet further experiments would be required to validate this assumption.

Shift experiments in agarose gels showed surprisingly clear bands, contradicting the somewhat accepted opinion of agarose as an inferior matrix for EMSAs (Berman et al. 1987; Hellman & Fried 2007).

3.4.3 Sepharose assays

Due to the initial struggles with assay quality in EMSA experiments and the lack of As(III) responsive unbinding, the Sepharose assay was developed. The results obtained with this assay were of relatively poor quality and the assay proved to be of little robustness when the typical inaccuracies of laboratory liquid handling equipment are regarded. That said, it did provide some data indicating that it could

be an alternative in the imaging of DNA-protein complexes. Perhaps the biggest advantage over the EMSA is the absence of a gel matrix, which can have multiple undesired effects on complexes. Electrophoresis further limits reliable detection of short lived complexes (3.4.2, Hellman & Fried 2007). The Sepharose assay is limited due to the necessity to image the protein without any staining. To make DNA localisation visible, probe DNA would have to be fluorescently labelled. Addition of competitor DNA in large quantities might reduce the binding capacity of Sepharose beads for specific DNA probes significantly. The use of longer DNA probes might increase the signal to noise ratio as each molecule would carry more negative charges, potentially binding more strongly to the Sepharose beads. In summary it can be said that the Sepharose assay might have some potential but requires considerable optimisation and characterisation before results could be regarded as reliable.

3.4.4 Generations of fusion proteins – to shift or not to shift

It appears that the initial cloning strategy was the root cause or at least involved in the lack of As(III) responsiveness observed in the EMSAs for ACH. While several other reasons might be involved, many can be excluded. The intense fluorescence of expression cultures, the intensifying bands on the SDS-PAGE gels with samples from the induction cultures and the results of the His-tag purifications all imply the presence of a complete fusion protein of correct mass. All gel shifts performed with these “first generation” proteins showed clear and specific shifting of probe DNAs once the essay quality had been enhanced. This leaves the question of whether it is an increased DNA affinity or a reduced As(III) reactivity that caused the seemingly permanent binding.

In either case, increased repression effects at a given As(III) concentration would be the consequence *in vivo* and interestingly similar effects of C-terminal fusions on ArsR function have been reported since the onset of this work. Merulla et al. unintentionally showed in 2013 that C-terminal mCherry fusions of ArsR via a peptide linker exhibit a greater stability of the DNA-protein complex due to unknown reasons.

Results presented in this work show that direct mCherry fusion to ArsR intensifies this phenomenon to a degree that DNA unbinding cannot be observed with the

employed assays, even at very high As(III) concentrations when compared with normal reaction levels. Thus it seems likely that the proximity of the fused protein component plays a crucial role in As(III) responsiveness or DNA affinity of the resulting chimeric protein. AGCH results show that increasing the distance between ArsR and mCherry by peptide linkers causes the fusion protein to regain some of the reactivity.

In agreement with the published findings of Merulla et al. and the assumptions described, reactivity is further increased by reducing the changes to the ArsR protein to the addition of a His-tag. This observation suggests that fusion of ArsR to various protein components might present a valuable tool for detection range tuning in biosensor applications.

The use of chimeric proteins following this strategy could also extend the practical range of sensors. This said, it remains unclear what possible influence the EMSA itself has on the As(III) responsiveness (3.4.2). Literature data concurs with the results obtained by the shift experiments in this work.

The As(III) concentrations required for DNA-unbinding in EMSAs assays are dramatically higher than in many other assays. When first characterised by Xu et al. in 1996, ArsR was reported to react to As(III) as sodium arsenite in a gel shift at concentrations around 5 mM. Contrary to these findings, Kawakami et al. later reported similar de-shifting to begin around 66 µM of As(III) as an undisclosed salt. Increasing this concentration 10-fold was sufficient to almost completely de-shift the assay (Kawakami et al. 2010). The latter results were obtained with ArsR-GFP fusion proteins, hence the expectation here would be to observe higher, not lower As(III) levels required.

The results presented in this work meet this expectation and seem to form a middle ground between the two extremes cited. ArsR-His showed initial unbinding of DNA in an EMSA at 50 µM As(III) with total unbinding occurring at 500 µM and above. To trigger the same reaction with AGCH proteins, concentrations needed to be increased approximately 10-fold, matching the numbers quoted by Xu et al. (1996).

Again, it remains unclear what influence the EMSA assay itself has on these results and Kawakami et al. did not publish a detailed protocol of the essays used in their study (Kawakami et al. 2010). The buffers used in all but the first EMSA

experiments in this work are based on the original publication by Xu *et al.* and thus should be adequate (Xu et al. 1996).

A potentially big influence is likely to be the DNA probe. Xu et al. used a relatively long DNA fragment in their experiments. The same group that had published the Kawakami results described a variation of their original sensor one year later (Siddiki et al. 2011). This time they reported that shortening the binding fragment and shifting it downstream to only incorporate one of the two binding sites greatly enhanced As(III) reactivity but also reduced overall DNA binding under arsenic free conditions. The 30 bp oligomer used in that study and the 50-mer oligo from the previous Kawakami publication have been compared to the oligos used in this work in figure 3.4-1. It becomes apparent that the binding sites as stated by the later publication are not entirely covered by the ArsR-BS oligos or the 50-mer oligo of the Kawakami publication. This might be the main reason for the increased As(III) responsiveness in the EMSAs when compared to the data published by Xu et al. in 1996, based on the longer binding DNA (figure 3.4-1).

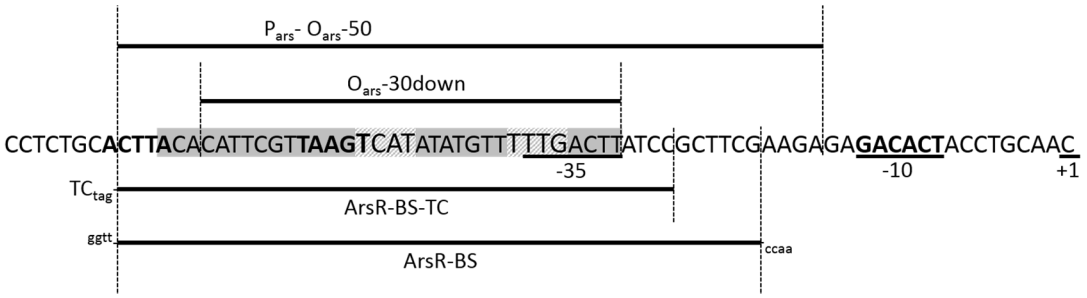

Figure 3.4-1: Binding sites and fragments of P_{ars}.

Binding sites originally published by Xu et al. are hatched, binding sites presented by Siddiki et al. are printed in bold. Promoter elements have been underlined and labelled accordingly. Shown above the nucleotide sequence is the coverage of the binding oligo employed by Kawakami et al. (P_{ars}-O_{ars}-50) and Siddiki et al. (O_{ars}-30down). Shown below are the binding oligos used in this work. Tags and overhangs are shown at their respective positions (Xu et al. 1996; Kawakami et al. 2010; Siddiki et al. 2011).

To show the specifity of ArsR DNA-binding activity, various competing probes were included in the shift assays performed. Widely accepted as the standard competitor in EMSA setups is salmon sperm DNA (Xu et al. 1996; Hellman & Fried 2007; Kawakami et al. 2010), although it remains unclear why sheared eukaryotic DNA is

thought to be an “all-time adequate” competitor for specifity studies of prokaryotic regulators. In addition to showing the specifity with salmon sperm DNA, the results presented in this work show no influence of P_{zntA} or a 40 bp fragment of its binding region on ArsR and perhaps still more important, in the presence of genomic *E. coli* DNA in the shifts with crude lysates. P_{zntA} carries an inverted repeat in its operator binding region and as such structurally resembles other metal inducible promoters, like P_{ars}. Based on these observations, unspecific binding of DNA by ArsR-His, ACH and AGCH can be precluded beyond reasonable doubt. Additionally, due to the design strategy of the AsGard sensor, such unspecific binding would not have been an issue as long as As(III) induced unbinding of DNA had still occurred.

Summarising the above discussion, the following conclusions based on the EMSA results can be made: all fusion proteins and ArsR-His show specific binding activity of the operator region of P_{ars}, yet only ArsR-His and AGCH show an As(III) responsive release. This release occurs in 10x lower concentrations for ArsR-His when compared to AGCH, likely as a consequence of protein fusion. The lack of response of ACH might be an extreme of this effect and it remains unclear if binding is truly unresponsive to As(III) or if further increased As(III) concentrations would lead to dissociation. Once these effects have been understood in greater detail, varying fusion partner size and proximity as well as probe DNA characteristics could represent a potent toolset for the engineering of powerful ArsR based As(III) sensors without the need for extensive protein mutation.

3.4.5 Proteins and surfaces, a sticky affair

At the core of the original AsGard plan stood a porous carrier material. For reasons of sensor cost and biodegradability this would have preferably been paper. Unfortunately, the results presented in this work indicate that the ArsR fusion protein AGCH has a strong tendency to bind a wide selection of surfaces, in most cases with surprising resistance to being washed off even after long washing steps and when the materials had been blocked by common blocking agents.

Binding to paper is somewhat understandable, as the large pores, uneven structure and unknown characteristics present a wide range of possible causes for protein retention. As such, the results obtained with the Hybond N+ membranes were

surprising. These positively charged nylon membranes with pore sizes of 0.45 µm are routinely employed for nucleic acid binding and are not efficient protein binders. They had been chosen as DNA immobilisation on them would have been ideal. Blocking by skimmed milk, BSA or Denhardt's solution would be expected to lead to very little attachment of AGCH added to the membranes afterwards. Yet, AGCH not only bound to the membranes but was impossible to wash off completely. Changing to a polyvinylidene difluoride (PVDF) membrane, the binding was enhanced. In contrast to nylon, this material does not carry strong surface charges and is supposed to have high protein retention abilities. Pores are less than half the diameter at 0.2 µm.

The similar behaviour on both membranes indicates that the attachment of AGCH is not charge related, excluding ionic interactions as a cause of binding. Repeating the experiment with nitrocellulose membranes showed a slightly reduced attachment, yet still impractical for application in the sensor. Nitrocellulose has an even smaller pore size than PVDF (0.1 µm) and lower protein retention characteristics. It remains unclear why the fusion proteins bound so readily to all porous materials.

EMSA experiments had shown that only a fraction of the loaded protein and DNA are involved in the formation of the shifted band representing the protein-DNA complex. As such it is possible that the purified protein samples contained a proportion of misfolded or otherwise impeded protein, readily binding to the membrane material. An influence of pore size combined with other factors appears likely, though it would be premature to draw this conclusion without further experiments. For reasons of time, it was decided to move on to non-porous materials. Interestingly, the experiments involving sanded PMMA slides indicate that AGCH has a tendency to bind to even this surface once it has been roughened by sanding.

Further experiments are necessary to elucidate the root cause of this binding tendency. In an attempt to utilise the binding effect, the sensor principle could be inverted, by labelling and detecting the DNA instead of, or additionally to the protein. Functionality of such a sensor would be dependent on the DNA binding capabilities of AGCH or any other ArsR-protein being unimpeded by surface binding.

3.4.6 PMMA – a tale of dots and spots

The unexpected behaviour of AGCH on porous surfaces in combination with the need for reliable DNA immobilisation led to the consideration of PMMA. It has been shown that the binding of TC-tagged DNA fragments to PMMA as described by Sun et al. (2012) can be adapted for increased volumes. A major concern was the effect of UV light on the DNA probes and thus the indirect influence on AGCH binding. The results presented allow for some general conclusions to be made.

a) dsDNA with a single strand TC overhang can be immobilised with the same process described for ssDNA immobilisation.
b) Immobilisation on the slide does not make DNA unavailable as a binding partner for AGCH.
c) UV radiation does not lead to DNA damage severe enough to prevent AGCH binding.

All of the above points had been concerns when the experimental setup was planned. It would have been possible that DNA would have become so tightly bound to the slides or damaged by UV light that AGCH would have been unable to bind. While this apparently did not occur in at least a fraction of the bound DNA molecules, there are some questions raised by the donut or coffee ring shaped fluorescence of AGCH after incubation and washing. Sun et al. have discussed the influence of surfactants, such as Triton, on the distribution of DNA probes (Sun et al. 2012). Their observations were that application of probes without any surfactant led to a shrinkage of the dots, ultimately resulting in a small spike of DNA in their initial center. Using too high concentrations of surfactant led to a concentration of DNA in a ring shape, similar to the localisation observed in this work.

The Sun et al. publication is based on 100 pl spots. Bearing in mind the greatly increased volumes of the AsGard dots, the concentration of Triton might have to be adapted to optimise DNA distribution. However, without analysis of the actual DNA distribution, for example by laser scanning, no statement can be made as to whether it is the DNA distribution that causes the ring appearance of AGCH fluorescence. Gel Green staining of the immobilised DNA did not show any evidence for greatly varying amounts of DNA within each spot. It is likely though that any irregularity would have to be quite extreme to be quantifiable by optical measurement of Gel

Green fluorescence distribution in this setup. Apart from the DNA distribution it is also possible that AGCH itself does not attach evenly to an otherwise evenly spread probe. Reasons for this could be the different availability for binding of the probe in different locations within each spot as an artefact of UV immobilisation or the concentration of AGCH at the outer rim of each droplet in consequence of evaporation during incubation. Interestingly, when sugars where added to increase drying resistance, fluorescence was no longer seen to concentrate in a ring shape.

It remains unclear whether this is due to changed contact angles of the protein solution on the slides or the sugars forming a layer during partial drying. Since samples containing sugars showed very little binding of AGCH to the DNA probes after washing, the explanation could equally be a dramatically reduced binding rate after application. Ideally these effects would be investigated in greater detail by changing the application patterns of AGCH or incubating the slides in an AGCH bath, rather than application of protein onto the slides. While these experiments are desirable, they would require larger amounts of purified protein than were available during this work. Another approach would be the addition of various surfactants to the protein solution. It was for this reason that Tween was added to some samples in the previously mentioned sugar experiment. Although these results are preliminary and experiments need to be repeated before final conclusions can be drawn, the data presented suggest that Tween does indeed reduce ring formation after washing, but not after incubation. In addition, fluorescence spots appear to be slightly smaller.

3.4.7 AsGard

As previously discussed, the binding of AGCH proteins to ArsR-BS-TC probes on PMMA slides was not without its flaws. Despite these limitations, As(III) in low concentrations was found to increase AGCH dissipation enough to make the effects visible with the imaging technique employed. It was shown that the presence of 0.134 µM As(III), corresponding to 10 ppb, leads to a clear reduction in remaining fluorescence after incubation for 5 or 10 minutes. Reduction of the same intensity has not been detected for metal free washing or experiments with As(V), indicating a specifity of the reaction. Employing image analysis software, the detection threshold for As(III) can be lowered further and even concentrations as low as 0.067 µM

As(III) have been shown to lead to a surprisingly clear reduction of fluorescence. The relatively narrow response window of the sensor does not allow for quantitative measurement at higher As(III) concentrations and relatively small differences in fluorescence make quantification in the lower ranges imprecise.

However, with the aim of a "drink or don't drink" feedback in mind, the behaviour of the AsGard sensor could be regarded as favourable, if it were not for the overall weak signal. One aspect of this problem is the chosen reporter, as detailed previously (3.4.1). Another aspect is the imaging by digital camera. This had been chosen in pursuing the aim of low-tech sensing but in the current AsGard setup it imposes limitations in precision and perhaps sensitivity, mainly because of background light from the SafeView imager.

Further practical issues are the struggle to obtain sufficiently focused pictures in the extreme dark. The long exposure pictures shown throughout the work might suggest a reasonable brightness, but even after long adjustment periods, no fluorescence was visible by eye after washing the slides. Initial trials to increase this fluorescence by giving the PMMA surface some texture failed and instead reduced assay quality.

There is a multitude of factors influencing the AsGard performance. Firstly, we must consider the amount of bound DNA probes and their availability for AGCH binding. It is likely that an unknown proportion of the DNA probes are either mutated or bound so tightly to the PMMA surface, that protein binding is inhibited. Secondly, the share of the AGCH proteins that can participate in DNA binding might be reduced due to misfolding or damage incurred during purification and storage, as indicated by EMSA. Lastly, the affinity of AGCH to the DNA probe may be poor. ArsR is known to have a relatively low affinity for its binding sites on the DNA and a small change of affinity upon metal binding when compared with other proteins, such as the MerR family metalloregulators (Merulla et al. 2013).

In previous biosensor setups, binding, washing and sensing have been performed in small buffer volumes. The sensor described by Kawakami *et al.* and modified by Siddiki *et al.* for example employed standard multi-well plates and measured either fluorescence retention or dissipation (Kawakami et al. 2010; Siddiki et al. 2011). In the small volumes used in such plates, ArsR or a fusion protein containing ArsR that has been released from the DNA probes will be present in a relatively high

concentration in the buffer. Consequently it will have a higher chance of coming into the vicinity of a DNA probe during the washing and sensing steps, giving it multiple chances of re-binding a probe.

In contrast, the 10 mL volumes employed during the corresponding steps in this project in conjunction with the small protein amounts, mean that any protein that releases the DNA is more likely to be "lost" in the solution, making re-binding occur less often. As previously mentioned, this might be an explanation for the constant depletion of fluorescence, even in metal free buffers, and in part for the relatively low detection limits.

To gain a deeper understanding of this effect, it would be desirable to vary the volumes employed. Moving to a plate reader friendly format could prove the most rewarding short-term experiment. Cutting the PMMA material into discs, fitting into the wells of a multiwell plate, would allow for the direct quantification of results. Some plate readers allow for the creation of heat maps of fluorescence by taking matrix measurements of each well. This would allow further studies of the "donut shape" and if combined with fluorescently tagged DNA probes, allow discrimination between uneven DNA immobilisation and uneven protein attachment. Ultimately, such aspects as real water sample sensing and the detection of different As species would become relevant as well, possibly addressable through the use of chemical reducing agents such as hydrogen sulfite (Rochette et al. 2000) or enzymes such as arsenate reductase.

3.5 Conclusion

While requiring considerable future investigation and optimisation, a foundation has been laid with this work for the creation of a super low-tech arsenic sensor, AsGard. In its minimalistic design, the AsGard sensor pushes beyond the widely accepted definitions of biosensors. It has been shown that the C-terminal fusion of mCherry as a reporter protein to ArsR leads to a change in As(III) responsiveness of ArsR. This change seems to be correlated to the distance between the two original protein sequences and corresponds to data found in literature. It remains unclear, whether too close fusion leads to a total loss or an extreme reduction of As(III) responsiveness. Functional fusion proteins of ArsR can be employed in a solid surface type sensor based on PMMA slides. Fluorescence from such fusion proteins initially co-localises with immobilised DNA probes but fades within minutes after the exposure to As(III) in concentrations corresponding to WHO limits for Arsenic in drinking water. As such the AsGard sensor would present a practical sensing concept if visibility, stability of the response and drying resistance can be enhanced.

4 Cell free heavy metal sensing

4.1 Introduction

The regular monitoring of drinking water presents a growing task in many industrialised and developing nations. Heavy metals originating from industrial pollution and ageing pipe networks are particularly challenging to measure. Conventional methods require expensive and bulky machinery, prohibiting on-site testing. In the past, the entailing requirements for centralised sample assessment and the resulting logistics have delayed the necessary counteractions or information campaigns. A recent, dramatic example, which reveals another factor, is the lead contamination of drinking water in the US city of Flint. Ill-considered decisions of local council members have led to large scale corrosive damaging of the lead pipes in the city's drinking water network. The resulting de facto poisoning of the drinking water affected over 100,000 people (Milman 2016; Olson et al. 2016). This and other examples show how centralised water testing led to delays and also present a single point of "failure" in this important health assessment (Olson et al. 2016). If the water testing was to be conducted in a decentralised way and, due to lower testing cost, at a higher frequency, then perhaps such cases could be avoided. In any case, alternatives to atomic absorption spectroscopy based heavy metal detection are desirable and many applications exist beyond the assessment of drinking water alone. Soil and water from industrial sites, quality assessment of processed sewage waters, irrigation water and private well testing are other fields of application where a decentralised testing routine could prove advantageous.

4.1.1 Aims of work presented in this chapter

Based on the aforementioned needs, a multiplexed biosensor based on metalloregulatory proteins becomes of interest. Each of these proteins is able to sense one or multiple metals (1.2.3) and in the right combination the response of a sensor comprising a selection of these regulators could become mutually exclusive. Such a sensor would need to be sensitive enough to detect heavy metals at relevant

concentrations and be free from legislative regulations prohibiting its use in the field. Hence, a cell free transcription translation based sensor appears to be a viable option to fulfil these requirements.

As such the aims of this chapter were:

- To construct 4 proof of concept, cell free transcription translation biosensors for the detection of heavy metals.
- To analyse the influence of heavy metals on a cell free expression system.
- To analyse the sensor sensitivities in a cell free expression system.
- To investigate the potential and characteristics of a tuneable sensor for As as a prototype for future biosensor designs.

4.2 Results

Initially, the Edinburgh arsenic biosensor (1.2.4.1.1) was modified to gain first experiences with the *in vitro* transcription-translation system. In its original design, the sensor consists of the *lacZ'α* and *arsR* genes under the control of P_{ars}, mimicking the feedback regulation system as it is present naturally in *E. coli*. Presence of As(III) in the cell or lack of ArsR will lead to the expression of both *arsR* and *lacZ'α* (figure 4.2-1). The latter gene encodes the α-fragment of β-galactosidase and when expressed, enables β-galactosidase activity in bacterial strains carrying the *lacZ*ΔM15 deletion by complementing this deletion. As a consequence, this system can only work if the bacterial strain carries the *lacZ* gene with the ΔM15 deletion in it. Any complete gene copy will lead to the uncontrolled expression of β-galactosidase. Due to this, it was necessary to determine whether the S30 extract system (Promega) employed in this part of the project carried an intact copy of the *lacZ* gene.

Figure 4.2-1: Plasmid map of the Edinburgh arsenic biosensor.

A 50 µL reaction as per the manufacturer's recommendations was incubated for 8 h at 37°C without any added DNA after the addition of 2.5 µM X-gal. The result was a clearly visible, blue colour, indicating β-galactosidase activity in the extract (figure 4.2-1). As a consequence of this result, the sensor construct had to rely on a reporter other than β-galactosidase to suit the requirements of the S30 extract system.

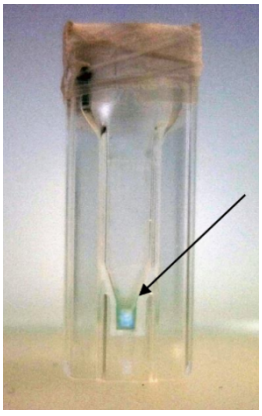

Figure 4.2-2: β-galactosidase activity in the S30 extract system. Activity is indicated by the blue colour (marked by arrow), which formed in the extract after 16 h incubation at 37°C with 2.5 µM X-gal.

4.2.1 Identifying suitable reporters

A pSB1A2 vector containing the aforementioned Edinburgh sensor construct (1.2.4.1.1 and figure 4.2-1) was amplified by PCR with specific PaperClip primers (2.2.1.5) binding outside the *lacZ'α* gene, pointing away from it. This allowed the amplification of the entire vector backbone with all previously described inserts except the *lacZ'α* gene, thus expressing ArsR and one of the tested reporters under the control of P_{ars} As reporter candidates, three genes were amplified from appropriate sources: the luciferase gene *luc,* the red fluorescent protein gene *mCherry* and a gene encoding for a β-glucosidase, which cleaves 4-methylumbelliferyl-β-D-glucuronide (MUG) resulting in fluorescence under UV light from 4-methylumbelliferone (MU). The latter gene originates from *Sorangium cellulosum* strain So ce56 and has the designated locus tag *sce4297*, abbreviated in this work to *sce*. The amplified genes were assembled with the backbone via PaperClip assembly and three biosensor plasmids were obtained, pSB1A2-*P_{ars}-arsR-luc*, pSB1A2-*P_{ars}-arsR-mCherry,* pSB1A2-*P_{ars}-arsR-sce* (figure 4.2-2).

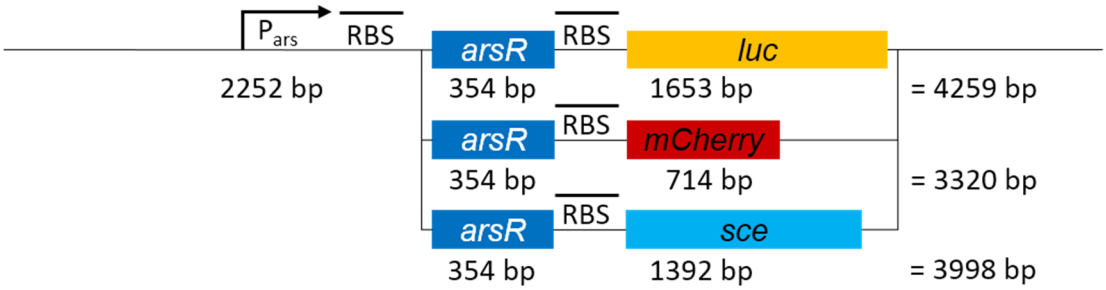

Figure 4.2-3: Cloning scheme of Edinburgh arsenic sensor based constructs.
Shown as a horizontal line traversing the figure is the pSB1A2 vector. Sequence elements originating from the original arsenic sensor construct are the P_{ars} promoter, *arsR* and both RBS. The three reporter genes employed are shown as coloured boxes in yellow, red and light blue and labelled accordingly. Relevant sequence lengths of each gene and the backbone including the intermediate RBS are given in bp below the corresponding parts and the total sequence length for each construct is given on the right. The diagram is not to scale.

These three sensor-constructs were tested in the S30 extract system according to the manufacturer's specifications. To ensure that the extract system itself was not influenced by As(III), the luciferase positive control provided with the system was performed with and without 1.335 µM As(III) present.

Results indicated that the luminescence was decreased by approximately 18% through the addition of As(III), although the variance between triplicates meant that this drop was far from being significant. Consecutive testing of all three plasmids in the S30 extract system with 0 µM, 0.067 µM, 0.200 µM, 0.667 µM and 1.335 µM As(III) revealed no As(III) dependent induction of any of the three reporters. However, the experiments showed that the luciferase assay was problematic to perform in any greater number. According to the manufacturer's documentation (Promega Corporation 2015a), the luminescence signal of this enhanced assay is nearly constant for about 20 s before it starts to decay with a half-life of 5 min. Given the foreseeable number of samples and the duration for each measurement, it was decided to drop luciferase as a reporter candidate. Consequently Sce and mCherry were chosen for further tests as reporters. The pBest plasmid from the positive control, carrying the IPTG inducible promoter-regulator P_{lac} was modified to include the *mCherry* or *sce* genes instead of the *luc* gene. Experiments with the S30 extract system revealed a comparatively low signal for *mCherry* and measurement results were so close to the detection limits of the employed Modulus multi mode reader that *sce* – which had given much stronger signals – was given the preference as a reporter.

4.2.2 Design and construction of metal responsive plasmids

Since the modified Edinburgh arsenic biosensor plasmids had performed poorly, a new approach had to be found. To reduce the translational demands, it was decided to remove the ArsR coding sequence and bring only the reporter under the control of metal responsive promoters. To ensure correct control of these promoters, the regulatory proteins were to be added separately to the S30 extract system. Consequently four metal responsive plasmids were to be designed to incorporate one of the following metal responsive promoters each: P_{ars}, P_{copA}, $P_{merTPAD}$ and P_{zntA}. As previously described (1.2.3), each of these promoters is controlled by a metal responsive protein and has been shown to induce expression in response to a selection of heavy metals. table 4.2-1 gives a brief overview of the regulatory proteins and the metal responsiveness of each promoter.

Table 4.2-1: Overview of metal responsive promoters, regulatory proteins and metal responsiveness.

Data collated from the available literature. See section 1.2.3 for details.

Promoter	Regulatory protein	Responsive to
$\mathbf{P_{ars}}$	ArsR	As, Sb
$\mathbf{P_{copA}}$	CueR	Ag, Au, Cu
$\mathbf{P_{merTPAD}}$	MerR	Hg
$\mathbf{P_{zntA}}$	ZntR	Cd, Hg, Pb, Zn

PaperClip assembly had been used successfully in the past and seemed ideal for the creation of a plasmid collection with the ability to extend this collection later without having to perform site mutagenesis to remove unwanted restriction sites. This assembly method leaves a GCC scar between the assembled parts. If this scar is localised between the RBS and the correlating gene, effects on transcription initiation can be detrimental, just as in traditional restriction-ligation cloning. In order to achieve maximum expression levels while keeping compatibility with the PaperClip system, some consideration was put into the optimisation of RBSs.

Protocols available for the self-fabrication of cell free expression systems are mainly based on cell extracts from the BL21 strain due to its deficiency in the Lon and OmpT proteases. The S30 expression system is described by the manufacturer as being deficient in these two proteases too. Additionally, the β-galactosidase activity (4.2.1) of the S30 extract hints towards the presence of an intact *lacZ* gene, as would be found in BL21. Under the assumption that the S30 extract supplied by Promega is based on the BL21 strain, predicted RBS efficiencies in this strain were simulated using the RBS Calculator (Salis et al. 2009; Espah Borujeni et al. 2014) in its most recent iteration (2.0) at the time.

The predicted translation initiation rates (TIR) of the original sequences were somewhat low (maximum 1 062 au) for the BL21 scenario. Changing the controlled genes to the *sce* and *mCherry* sequences with a GCC scar, led to a further drop in predicted TIR. Employing the design features of the RBS calculator, synthetic RBSs were designed for each promoter. Sequence constraints were altered to create synthetic sequences that incorporated the GCC scar as a part of the RBS. Sites of

various lengths ranging from 7 to 22 bp were computed and compared in their predicted TIR. While there was a general increase with greater sequence lengths, this increase was most prominent in short sites and less pronounced when the sequences passed a certain length.

Practical considerations of oligonucleotide availability and practicality in PCR put an emphasis on maximum TIR at minimal altered sequence length. Due to this, the sRBS sequences were chosen to represent an optimal activity/length ratio. The regulatory mechanisms of the employed promoters (except P_{ars}) was another aspect to be considered. For the MerR-family, the mechanism is based on spatial orientation of the DNA (1.2.3.2). Changing the promoter sequence too close to the regulatory region was anticipated to have detrimental effects on the inducibility of the promoters.

Consequently, synthetic RBS (sRBS) including the GCC scar of 15 bp for P_{ars}, 21 bp for P_{copA}, 14 bp for $P_{merTPAD}$ and 16 bp length for P_{zntA} with TIRs between 12 463 and 408 360 were designed (figure 4.2-3). In the latter three cases, these sites are well separated from the regulatory regions and the transcription start site. figure 4.2-4 shows an overview of the promoter sequences and their respective sRBS.

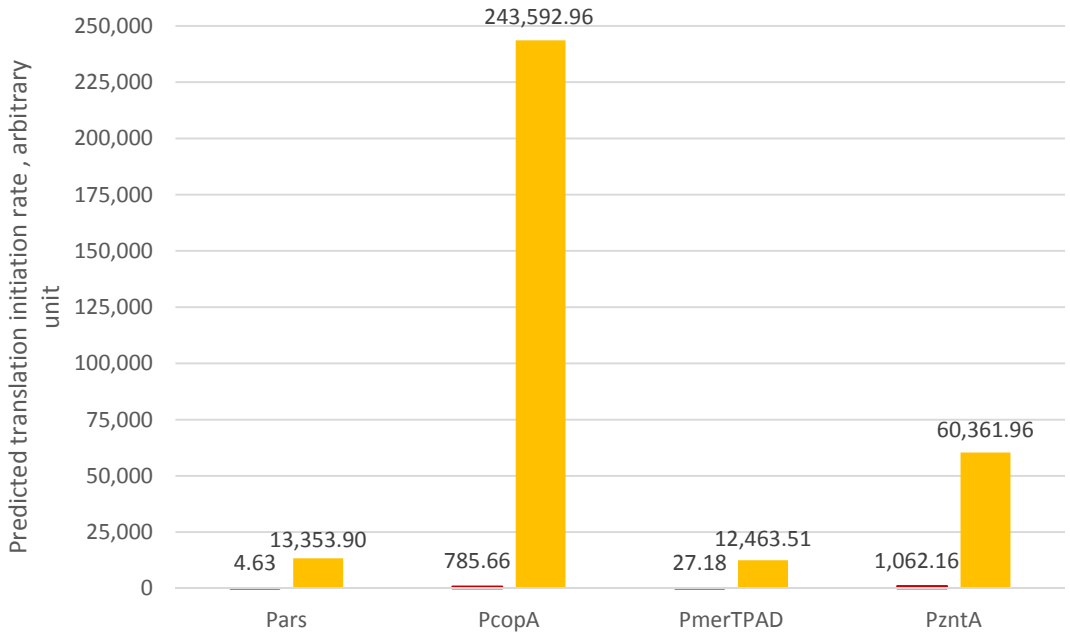

Figure 4.2-4: Predicted translation initiation rates for Sce under control of the metal inducible promoters.

Predictions were made with the RBS Calculator using the free energy model 2.0. Red bars (left in each data set) show the predicted TIR for the wild type RBS, orange bars (right in each data set) show the corresponding value for the sRBS.

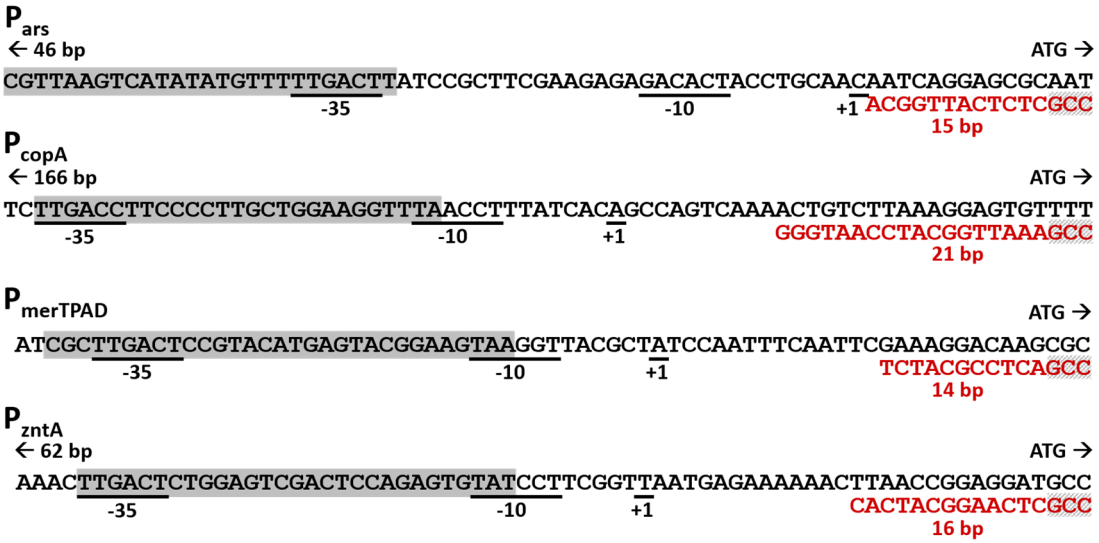

Figure 4.2-5: Partial promoter sequences of P_{ars}, P_{copA}, $P_{merTPAD}$ and P_{zntA}.

For $P_{merTPAD}$ the entire promoter sequence is shown while the remaining sequences have been abbreviated as shown above the corresponding sequence with a left pointing arrow (←). For all sequences, the areas of protein-DNA interaction published in the according literature are marked with a grey box. -35 and -10 promoter regions as well as the transcription start (+1) are underlined and labelled accordingly. Sequences of the synthetic RBS are shown underneath the native sequences in red. The GCC scar of the PaperClip assembly method is marked by a grey dashed box. All sequences end before the start codon of the controlled genes, marked by ATG and an arrow on the right (Xu et al. 1996; Brocklehurst et al. 1999; Outten et al. 2000; Brown et al. 2003; Siddiki et al. 2011).

The promoter sequences were amplified from genomic *E. coli* DNA and laboratory stocks. The Sce gene was amplified from a laboratory stock and proved to be a challenging sequence to work with. Considerable trouble-shooting and optimisation was necessary before the 67% GC gene had been combined successfully with all other components. Initially a total of 10 constructs was planned, incorporating each of the metal responsive promoters with an engineered RBS (sRBS) and the positive control plasmid containing the same reporter genes under the control of P_{lac} from the pBest vector.

As described previously, the positive control with *mCherry* and *sce* had revealed the vast superiority of the latter and cloning of *mCherry* constructs was stopped when this data became available. After *in vivo* studies presented in the next sections had shown little expression and limited response to metal ions for some of the

combinations, versions with a native RBS were cloned for all combinations, adding another 4 plasmids to the library. These clones were created by MABEL (2.2.1.7), using the sRBS plasmids as templates. All constructs were confirmed by sequencing. figure 4.2-5 gives an overview of the cloning and naming scheme used throughout this chapter.

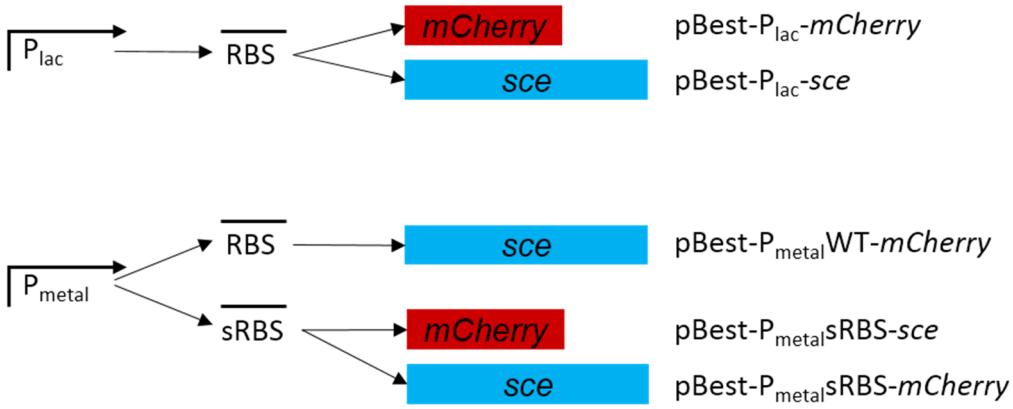

Figure 4.2-6: Cloning scheme of metal sensing plasmids based on the pBest vector.
The positive control plasmid, bearing the original *lac* promoter was combined with *mCherry* and *sce*. The metal inducible promoters, symbolised by P_{metal} in this figure, were combined with the same two reporter genes but each with the native and an engineered RBS (sRBS). The schematic naming strategy as used throughout this work is given on the right. Note that the *mCherry* combinations were later dropped and not all promoters were paired with *mCherry*.

4.2.3 Influence of heavy metals on MU-fluorescence in *E. coli*

To assess the influence of the heavy metals, *E. coli* JM109(DE3) cultures containing the positive control construct P_{lac}-*sce* were employed in an *in vivo* study. Fresh overnight liquid cultures were used to inoculate (1% inoculum) a main culture in LB with the appropriate antibiotics. Main cultures were then grown to an OD_{600} of 0.7, continuously shaking at 37°C. For assays in 96-well plates, each sample was individually supplied with MUG substrate to a total concentration of 0.2 mM, induced with 0.5 mM IPTG and various concentrations of heavy metals were added. To identify maximum tolerable metal levels in terms of growth, *sce* expression and activity, a selection of concentrations was tested for each metal. Effects of various metals differed widely.

Experiments with Au (figure 4.2-8) revealed a limited influence of up to 250 µM Au within the first 4 h. After this time point the fluorescence levels of all cultures containing Au began to show slower increases when compared to the metal free positive control. A concentration of 500 µM led to a similarly shaped curve but at generally reduced fluorescence levels. The highest concentration tested, 1000 µM, practically inhibited any fluorescence development. A concentration of 250 µM to 500 µM was deemed the maximum for later essays, keeping in mind that at 500 µM some negative influence was to be expected. OD_{600} measurements from the same experiments indicated that the discussed Au concentrations affected the growth of the cultures proportionally stronger than the fluorescence development, reflecting cell toxicity rather than inhibition of Sce.

Effects of Hg were similar (figure 4.2-9). All but the highest concentration (25 µM) had little effect up to 3.5 h, after which fluorescence increase began to slow down for the second highest concentration samples at 10 µM. Cell growth as measured by OD_{600} reflected this behaviour closely. With the precautions aforementioned for Au, maximum testing levels in future experiments for Hg were set at 10 µM

Ag showed relatively little influence at concentrations up to 10 µM, with a 25 µM concentration leading to a temporary stagnation in fluorescence development after 3 h (figure 4.2-10). Consequently 10 µM was set as the upper limit for future experiments. Effects of Ag were most prominent in the OD_{600} measurements, similarly to the observations made for Au.

Cd showed no influence on the assay at 10 µM, but increasing the concentration to 100 µM, 250 µM and 500 µM each led to a profound reduction in fluorescence increase (figure 4.2-11). Interestingly Cd seemed to have almost no influence on cell growth, indicating that Sce was being inhibited while cells grew almost normally. As such it was difficult to set a designated maximum concentration for the following experiments. Literature data reports induction of P_{zntA} *in vivo* at 100 µM Cd (Brocklehurst et al. 1999; Binet & Poole 2000). Consequently the maximum limit was set to this concentration.

Cu had relatively little effect, with each concentration increase leading to a slight drop in fluorescence increase up to a maximum tested concentration of 1500 µM (figure 4.2-12). OD_{600} measurements for these experiments indicated enhanced

growth with increasing Cu concentrations. However, as Cu in the tested concentrations can lead to the formation of a visible blue colour in the medium, it is likely that increade OD_{600} readings reflect this colour, rather than a truly increased bacterial growth. The maximum for future tests was set to 1000 µM to 1500 µM, depending on the experiment.

Adding up to 1000 µM Pb to the assay had a slightly enhancing effect on the fluorescence increase and OD_{600} development (figure 4.2-13). Consequently concentrations up to this mark were deemed adequate for future works.

Setting a limit for Zn proved difficult, similarly to the struggle with Cd in this regard. Even the lowest tested concentration of 500 µM had a profound effect, reducing fluorescence in these samples to little more than a third of the positive control after 3 h (figure 4.2-14). Similarly to Cu, OD_{600} data may indicate little influence of Zn on bacterial growth but the effects of increased turbidity due to resolved Zn in the medium should be taken into account. Zn is widely known as a potent antibacterial agent above certain concentrations, underlining this call for caution in interpreting the OD_{600} results. Literature research had indicated that strong induction was to be expected to overcome some of the toxic effects at concentrations between 250 and 1500 µM (Brocklehurst et al. 1999). 500 µM were accepted as a limit for negative controls in future experiments, but higher levels were to be tested in P_{zntA} induction experiments.

All metal limits as specified above are collated in figure 4.2-7 for easier comparison of their respective influences. When comparing the fluorescence development at the chosen concentrations it became apparent that Au, Cd, Cu and Zn would be expected to have a greater reducing effect when compared to the other metals. In this regard Zn showed the most pronounced influence. None of the chosen maximal concentrations reduced the final OD_{600} to less than approx. 70% of the metal free control cultures.

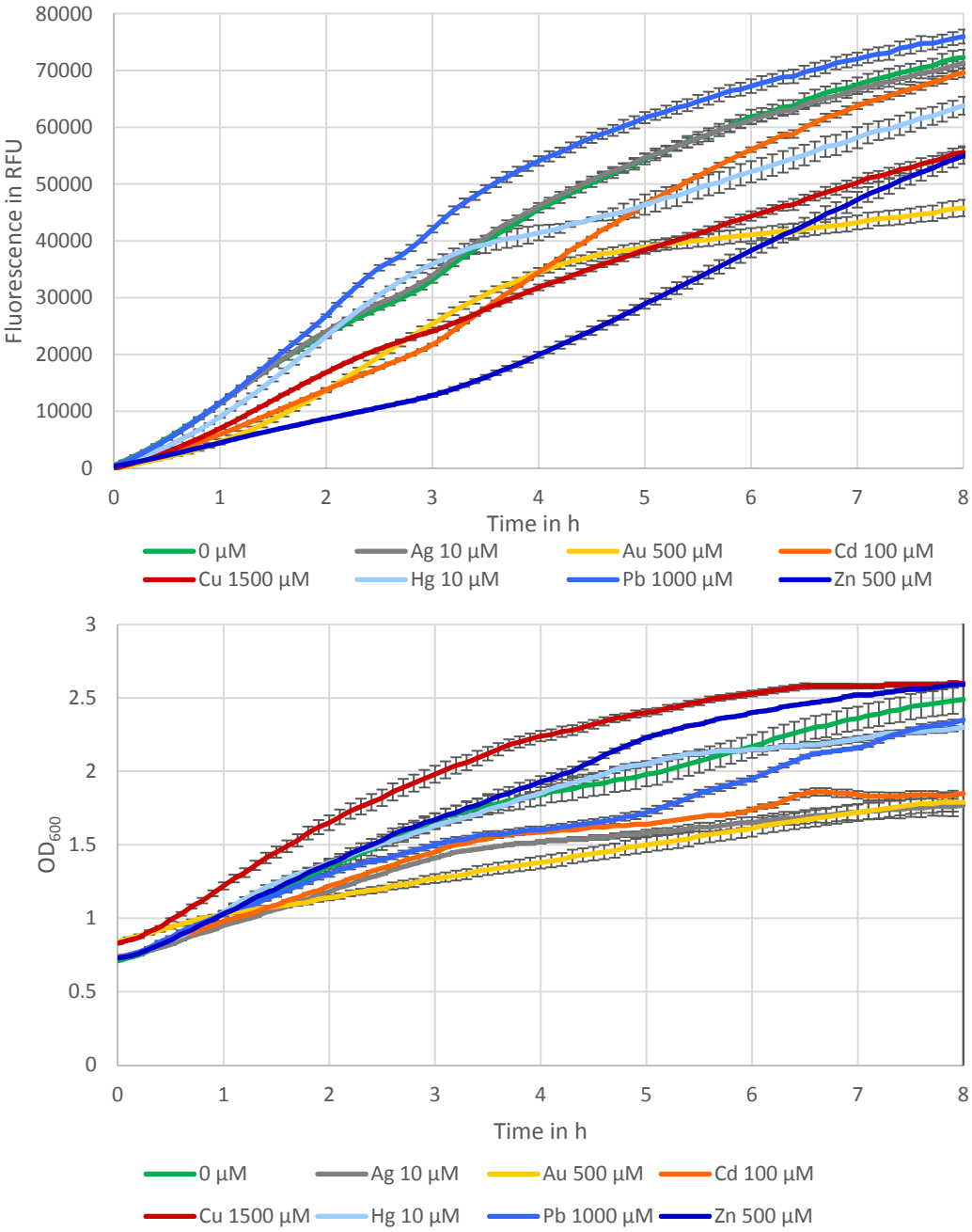

Figure 4.2-7: Influence of heavy metal chosen maximum controls on *E. coli* pBest-P_{lac}-*sce*. Cultures were grown to an OD_{600} of 0.7 prior to addition of 0.2 mM MUG, 0.5 mM IPTG and metal as specified in the legend. Top half of figure shows fluorescence measured at 355/460 nm. Graphs represent negative control (no plasmid) corrected average of triplicates. Bottom half of figure shows average of triplicate OD_{600} measurements. All error bars show one SD above and below.

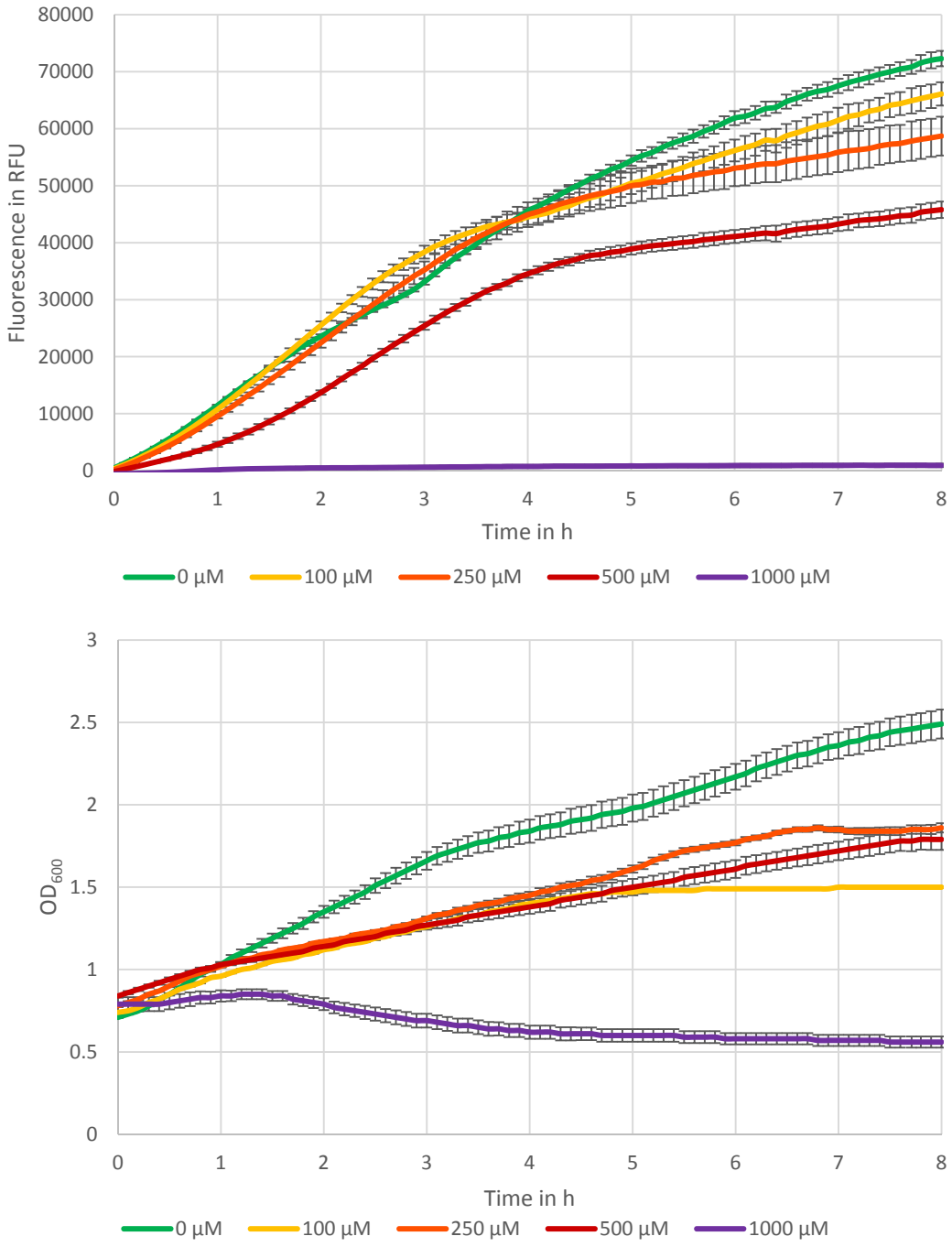

Figure 4.2-8: Influence of Au as $AuCl_3$ on *E. coli* pBest-P_{lac}-*sce*.
Cultures were grown to an OD_{600} of 0.7 prior to addition of 0.2 mM MUG, 0.5 mM IPTG and metal as specified in the legend. Top half of figure shows fluorescence measured at 355/460 nm. Graphs represent negative control (no plasmid) corrected average of triplicates. Bottom half of figure shows average of triplicate OD_{600} measurements. All error bars show one SD above and below.

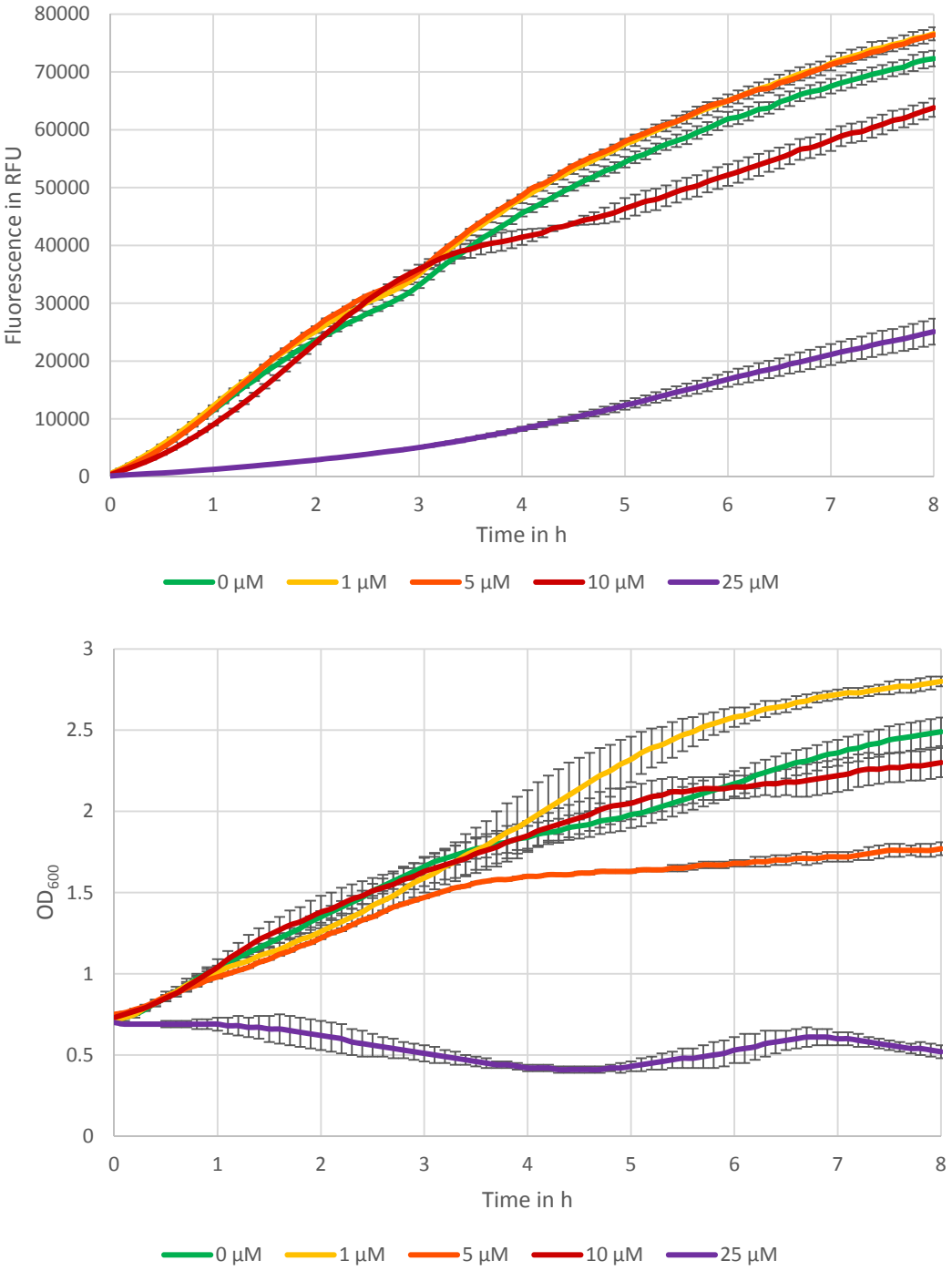

Figure 4.2-9: Influence of Hg as $HgCl_2$ on *E. coli* pBest-P_{lac}-*sce*.
Cultures were grown to an OD_{600} of 0.7 prior to addition of 0.2 mM MUG, 0.5 mM IPTG and metal as specified in the legend. Top half of figure shows fluorescence measured at 355/460 nm. Graphs represent negative control (no plasmid) corrected average of triplicates. Bottom half of figure shows average of triplicate OD_{600} measurements. All error bars show one SD above and below.

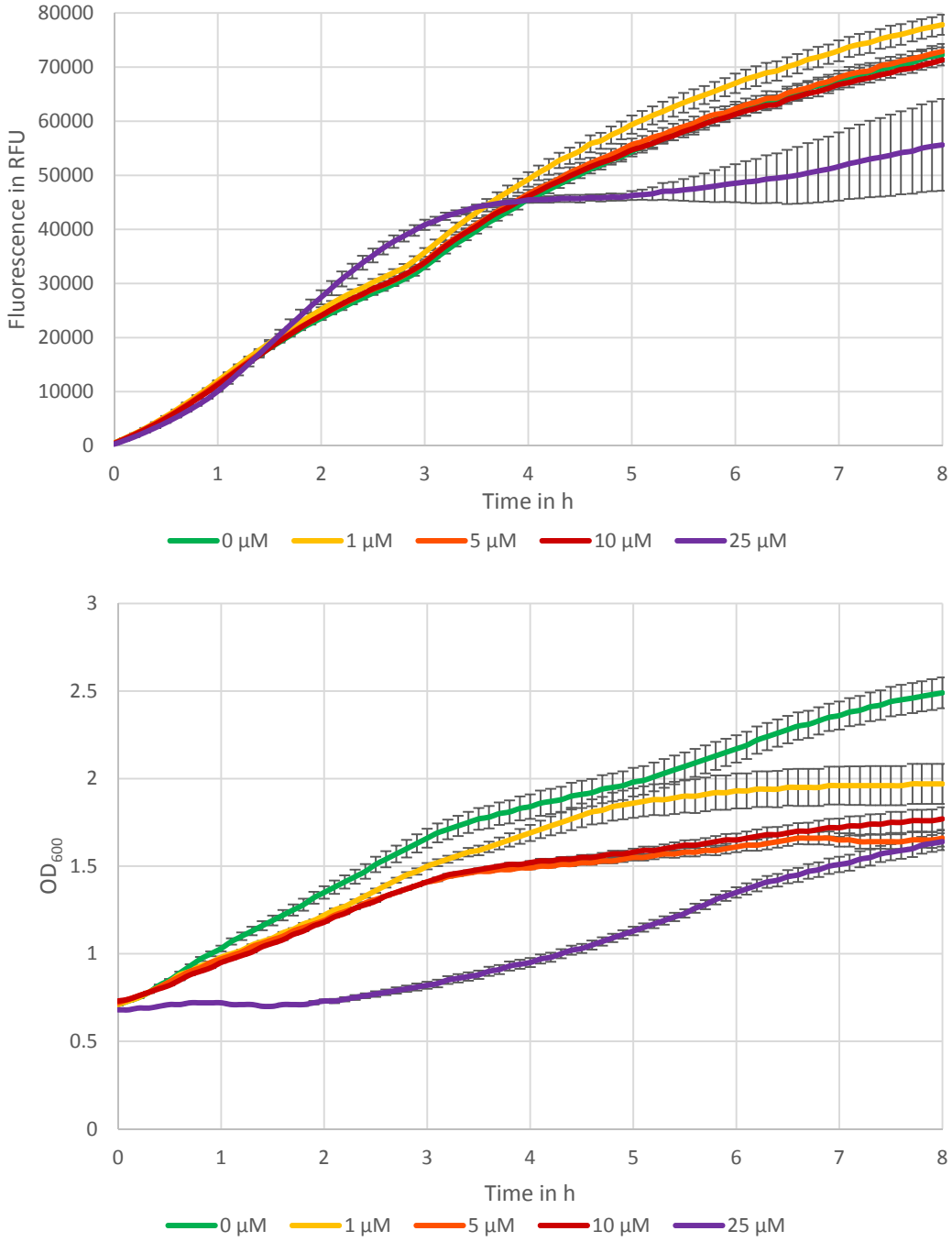

Figure 4.2-10: Influence of Ag as $AgNO_3$ on *E. coli* pBest-P_{lac}-*sce*.
Cultures were grown to an OD_{600} of 0.7 prior to addition of 0.2 mM MUG, 0.5 mM IPTG and metal as specified in the legend. Top half of figure shows fluorescence measured at 355/460 nm. Graphs represent negative control (no plasmid) corrected average of triplicates. Bottom half of figure shows average of triplicate OD_{600} measurements. All error bars show one SD above and below.

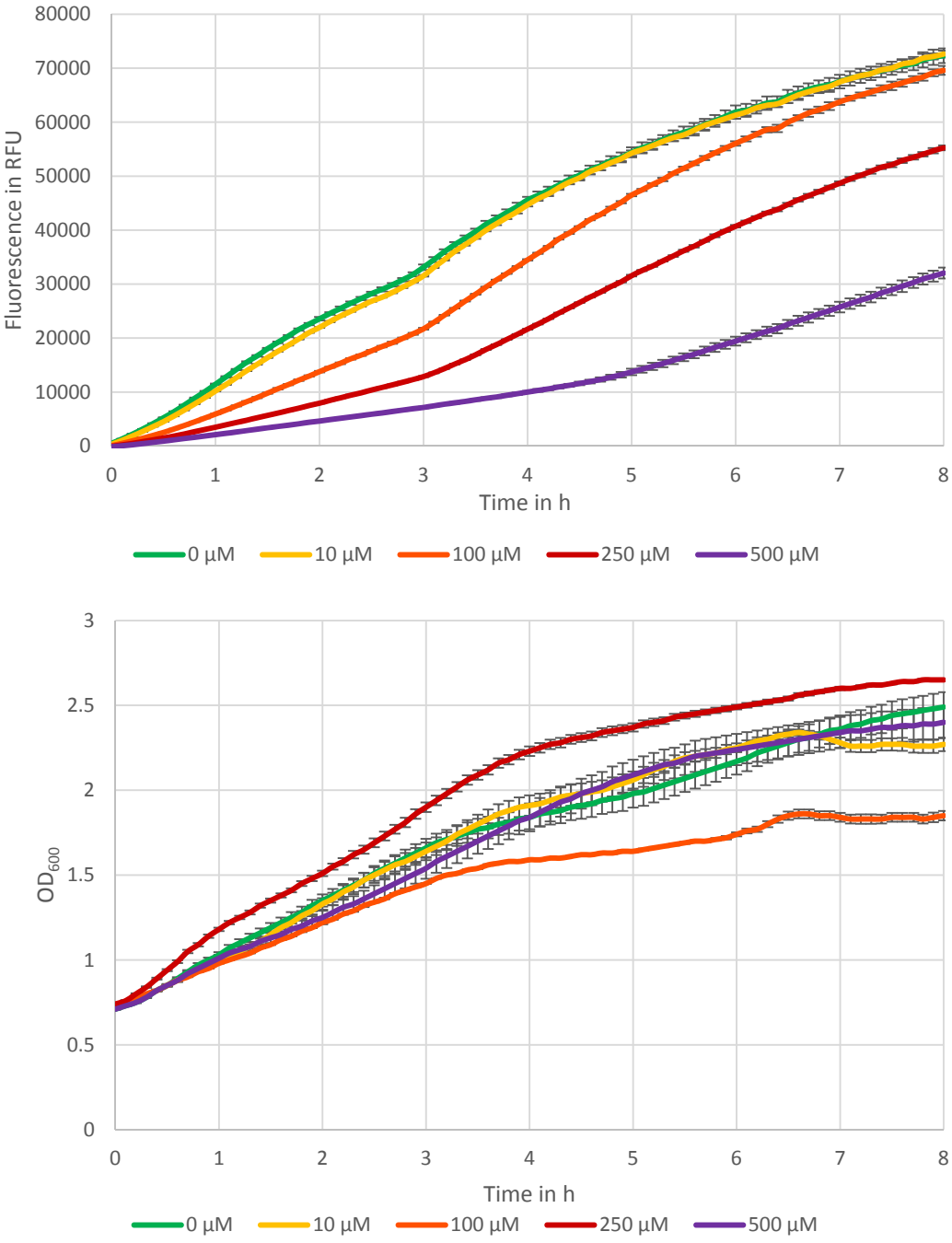

Figure 4.2-11: Influence of Cd as $CdCl_2$ on *E. coli* pBest-P_{lac}-*sce*.

Cultures were grown to an OD_{600} of 0.7 prior to addition of 0.2 mM MUG, 0.5 mM IPTG and metal as specified in the legend. Top half of figure shows fluorescence measured at 355/460 nm. Graphs represent negative control (no plasmid) corrected average of triplicates. Bottom half of figure shows average of triplicate OD_{600} measurements. All error bars show one SD above and below.

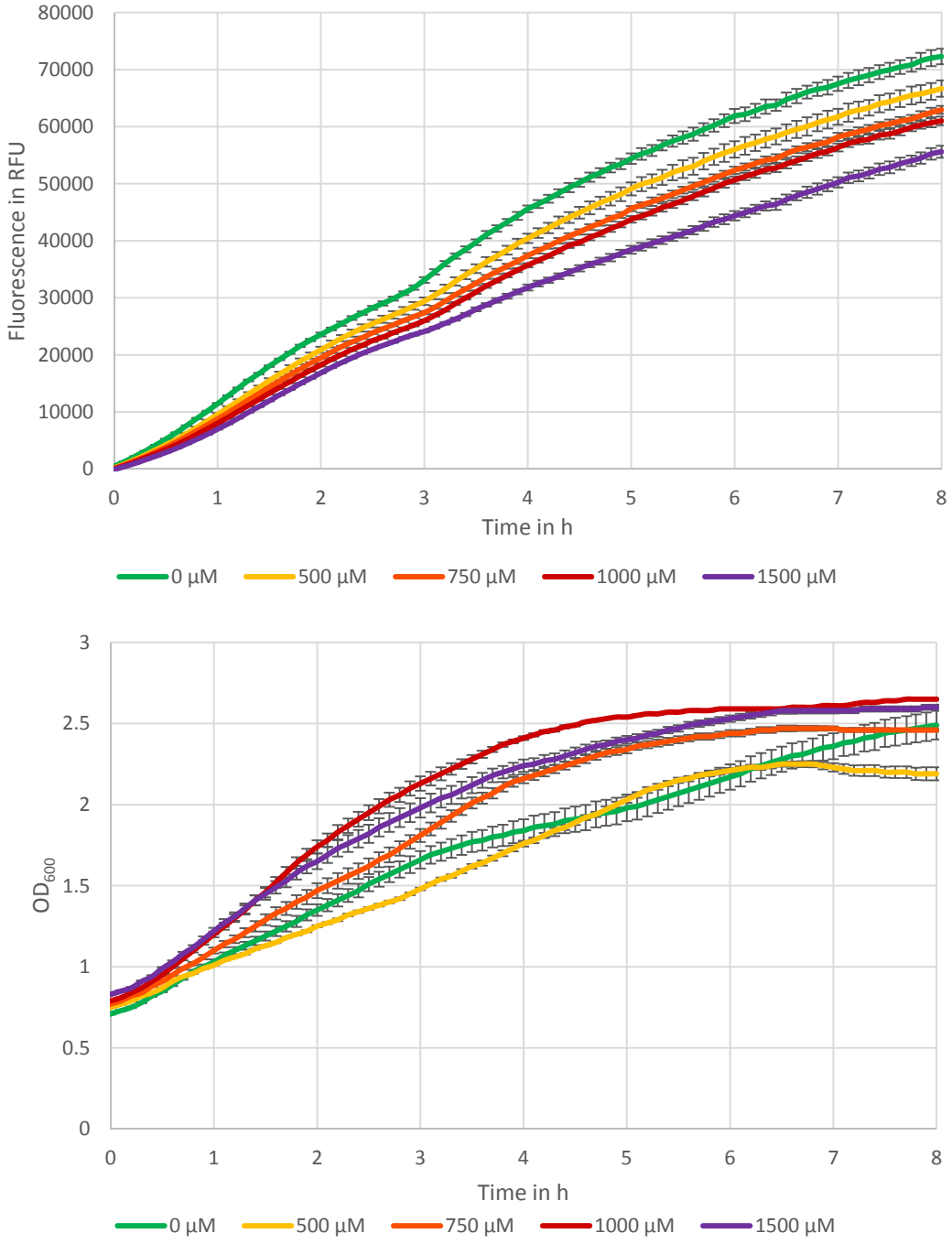

Figure 4.2-12: Influence of Cu as $CuSO_4$ on *E. coli* pBest-P_{lac}-*sce*.
Cultures were grown to an OD_{600} of 0.7 prior to addition of 0.2 mM MUG, 0.5 mM IPTG and metal as specified in the legend. Top half of figure shows fluorescence measured at 355/460 nm. Graphs represent negative control (no plasmid) corrected average of triplicates. Bottom half of figure shows average of triplicate OD_{600} measurements. All error bars show one SD above and below.

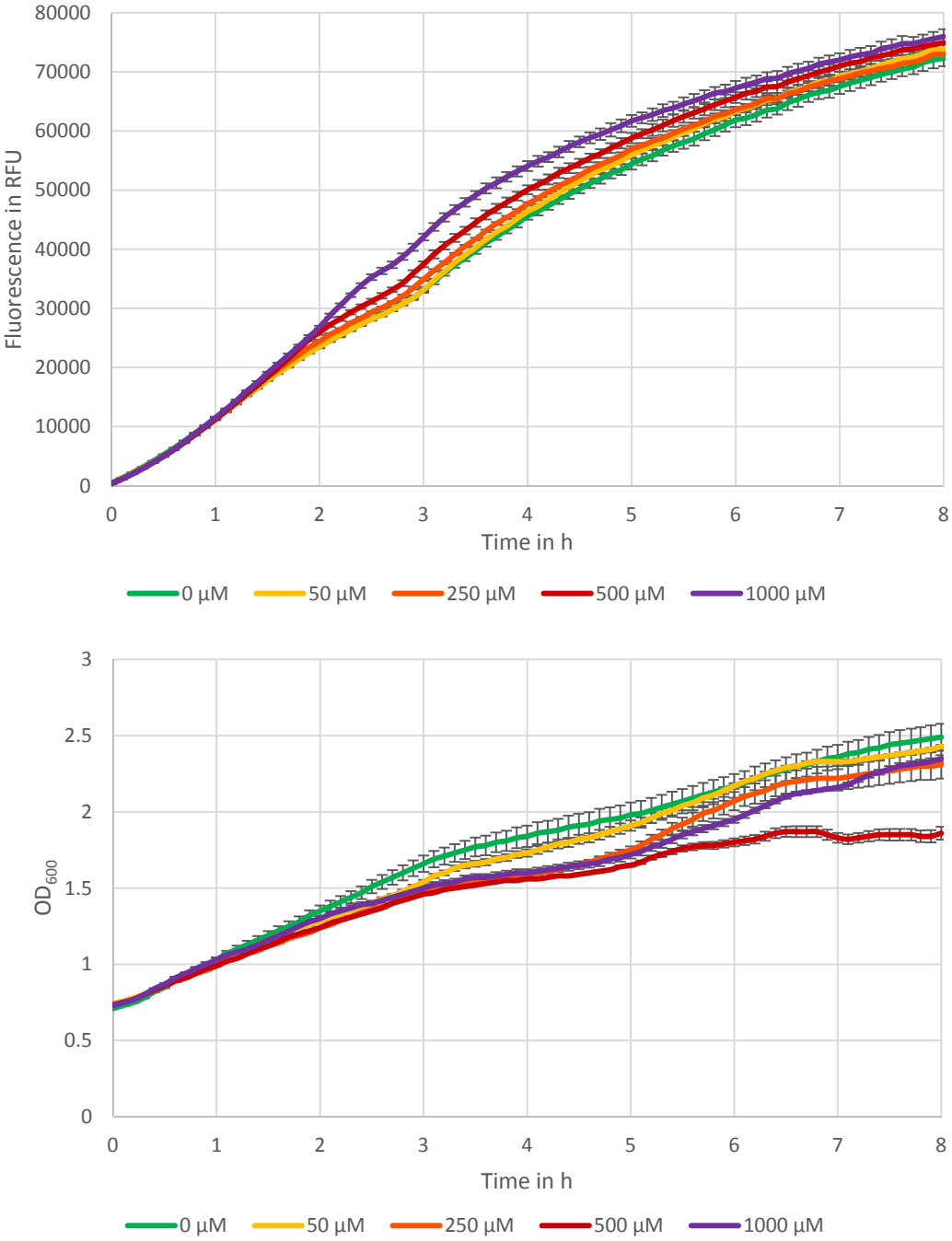

Figure 4.2-13: Influence of Pb as $PbNO_3$ on *E. coli* pBest-P_{lac}-*sce*.
Cultures were grown to an OD_{600} of 0.7 prior to addition of 0.2 mM MUG, 0.5 mM IPTG and metal as specified in the legend. Top half of figure shows fluorescence measured at 355/460 nm. Graphs represent negative control (no plasmid) corrected average of triplicates. Bottom half of figure shows average of triplicate OD_{600} measurements. All error bars show one SD above and below.

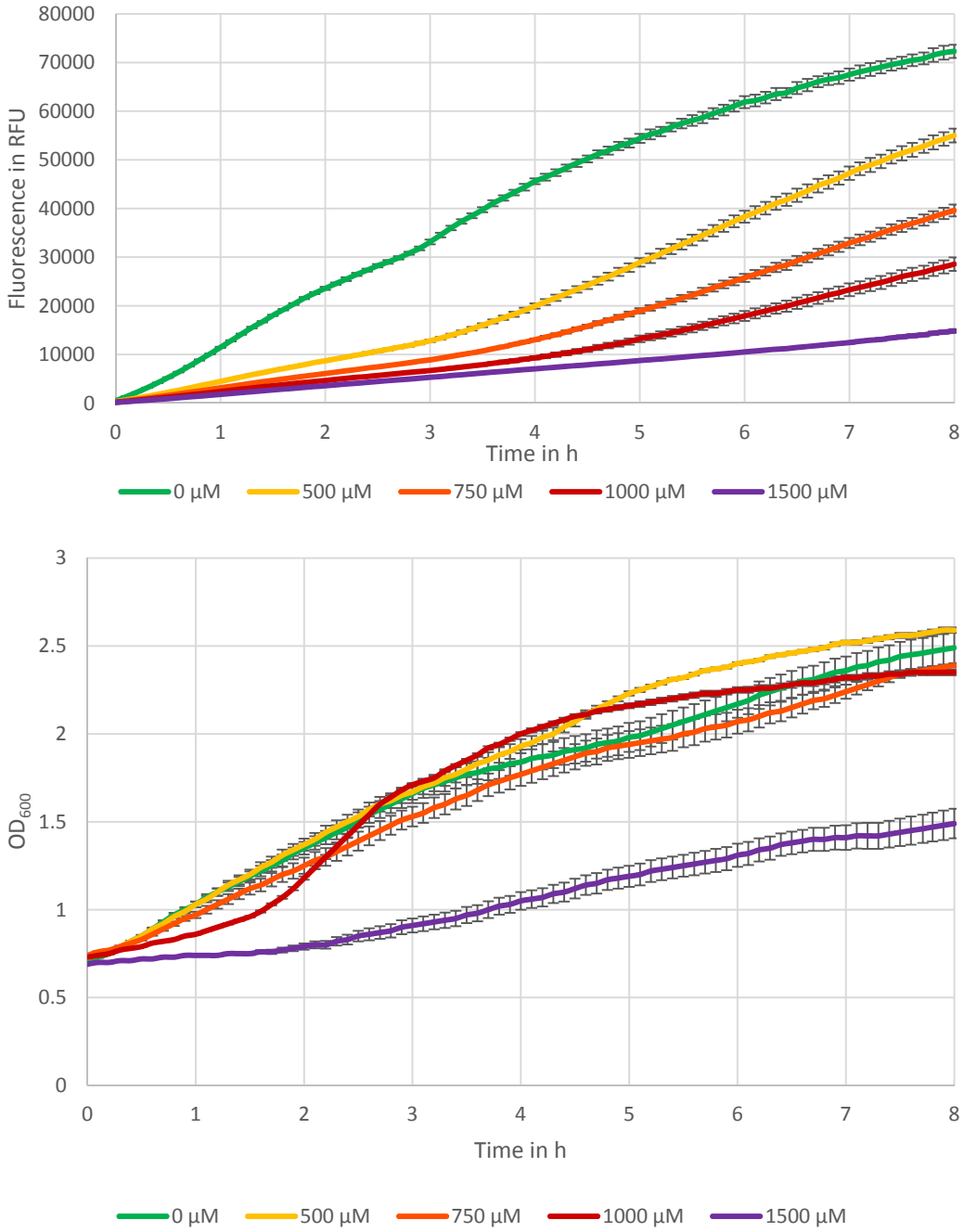

Figure 4.2-14: Influence of Zn as $ZnCl_2$ on *E. coli* pBest-P_{lac}-*sce*.
Cultures were grown to an OD_{600} of 0.7 prior to addition of 0.2 mM MUG, 0.5 mM IPTG and metal as specified in the legend. Top half of figure shows fluorescence measured at 355/460 nm. Graphs represent negative control (no plasmid) corrected average of triplicates. Bottom half of figure shows average of triplicate OD_{600} measurements. All error bars show one SD above and below.

4.2.4 Heavy metal induction of the metal sensing constructs *in vivo*

In vivo the metal responsive operons are largely based on single feedback loops. Within these loops the regulated promoter is controlling the expression of its own regulating protein. The constructs designed for cell free metal sensing did not incorporate this feedback loop. This meant that cells bearing the high copy number metal sensing plasmids would likely exhibit a severe imbalance in the regulatory circuit due to the limited amount of regulatory proteins supplied by the natively present operon copy in the cell. Since resources for cell free sensing experiments were limited, it was decided to perform a preliminary *in vivo* test of the metal responsive properties of each construct to preselect candidates for *in vitro* testing. Experiments were performed as previously detailed (4.2.3). Metal concentrations were chosen based on literature research and the results from *in vivo* testing. A positive control reaction with the P_{lac}-*sce* construct was performed together with each set of experiments to ensure consistency of the assays and as a reference point for background activity of the uninduced metal promoters.

Initially only the sRBS versions of the plasmids were available and first tests with the copper inducible promoter P_{copA} showed little induction in this constellation. Cu concentration of 1 µM to 1500 µM as well as Ag, Au, Zn, Cd, Hg and Pb did not show any clear change in fluorescence development when compared to the metal free control. Due to this, the native RBS versions of all constructs were cloned and added to the testing regimen. In the case of P_{copA} this led to a drastic change in metal responsiveness. All promoters exhibited basal activity. To make data comparable between test runs and with data found in the literature, raw data was processed to allow for the expression of results as a fold increase in Sce enzyme expression. Raw data (figure 4.2-15 A) was blank corrected and the noisy OD_{600} measurement curves were smoothed by exponential smoothing (figure 4.2-15 B). The second derivative was calculated from the fluorescence curves (figure 4.2-15 C) and values were divided by the smoothed OD_{600} data. The resulting curves exhibited a peak at the time point where the fluorescence had been increasing at the maximal rate (figure 4.2-15 D), corrected for cell growth and thus partly allowing for cell toxicity effects. The peak value for each sample was divided by the peak value of the uninduced

negative control, allowing for the expression of data as fold induction of basal activity.

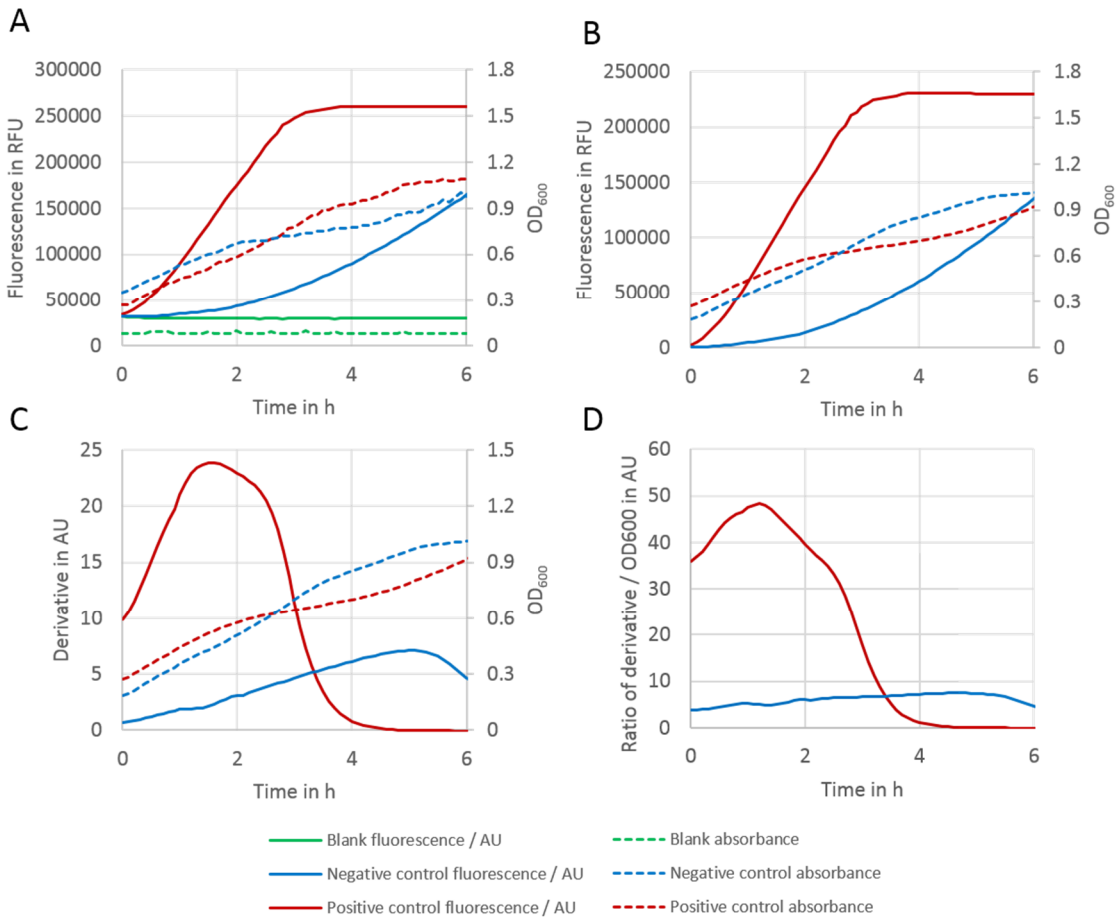

Figure 4.2-15: Data processing steps of fold induction calculations.
Shown is the raw data and subsequent calculation results for uninduced P_{zntA}-*sce* negative control and P_{lac}-*sce* positive control. A) Raw data for fluorescence measured at 355/460 nm and absorption at 600 nm. B) Blank corrected fluorescence measurements and blank corrected absorption data after exponential smoothing. C) Derivative of fluorescence data and absorption data as shown in B. D) Ratio of derivative and absorption.

To visualise this background activity and allow for a rough comparison in induction strength, the same calculations were performed for the positive control (P_{lac}-*sce*) in each experiment. The resulting numbers, although strictly speaking not fold induction of the same promoter (as they originated from a different promoter), indicated the relative background level. As the positive control expression level was stable, lower fold-induction values calculated for the positive control in relation to uninduced controls imply high background in the latter samples and vice versa.

Metal induction reaching the values obtained with the positive control consequently indicated high promoter activation in these samples.

The data obtained showed a lack of metal reactivity in the P_{copA}-sRBS construct with no sample reaching over 1.2-fold induction. Background activity of this promoter version was low compared to the other promoters including the unmodified P_{copA}. The expression of the positive control was approximately 7-fold that of the uninduced control.

The P_{copA} promoter with WT RBS showed a Cu response. Low concentrations from around 5 µM to 250 µM led to a 1.4-fold to 2-fold induction. Increased concentrations up to 1500 µM led to a 3.5-fold increase in expression, slightly less than half of the 7.8-fold induction of the positive control (figure 4.2-14).

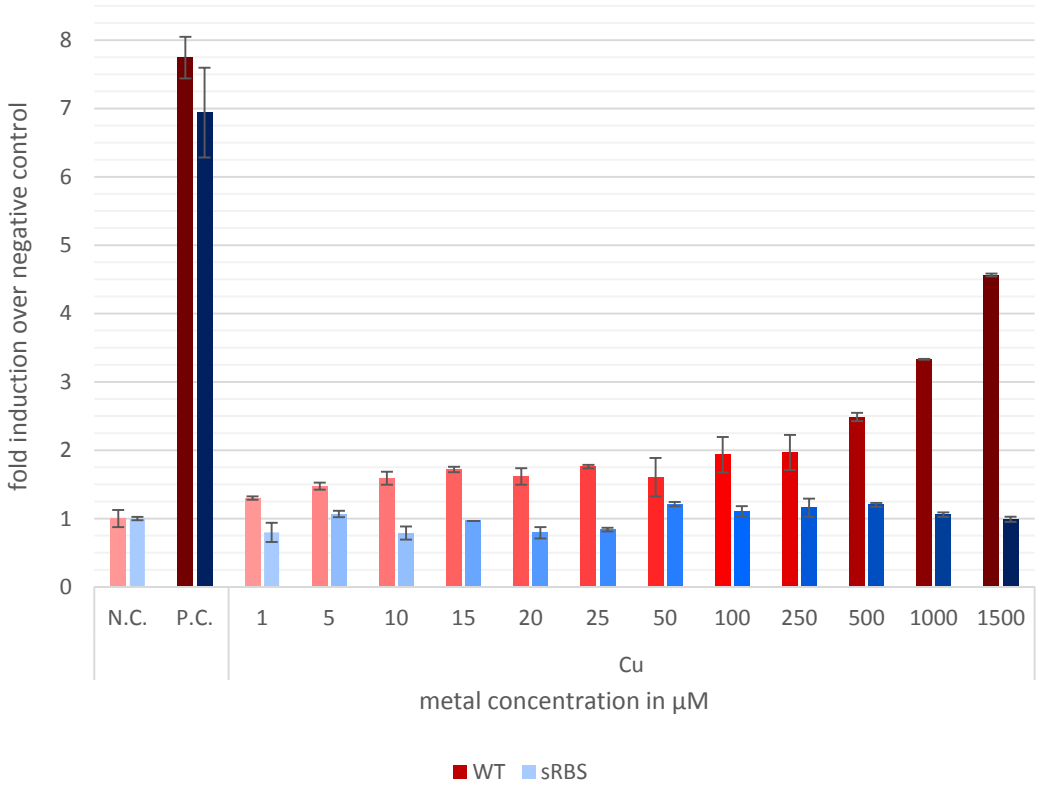

Figure 4.2-16: Cu induction of P_{copA} *in vivo*.
Induction of *sce* expression is shown as fold induction of the metal free negative control. P_{lac} control (P.C.) is shown as benchmark for maximum activity and background activity of the uninduced control (N.C.). Error bars show one SD above and below the average of triplicates.

Literature reports the induction of P_{copA} by Ag and Au (Stoyanov et al. 2001; Stoyanov & Brown 2003; Stoyanov et al. 2003). Ag induction is reported to be stronger and more sensitive than Cu and Au induction. Addition of 5 µM Ag to the assay led to a 1.8-fold increase, which had developed from no induction at 2 µM and below. Au was the only metal tested in a concentration series that showed a reduction at high concentrations after having previously led to very strong induction, indicating the onset of toxic effects. Expressed numerically, increasing the Au concentration from 10 µM to 100 µM led to an 8-fold increase in expression, more than the positive control at 7.8-fold. Further increase of Au concentration to 500 µM reduced the expression level back to 2-fold.

Zn and Cd did not lead to any increase in expression. Hg and Pb led to weak increases of 1.2-fold and 1.8-fold respectively (figure 4.2-15).

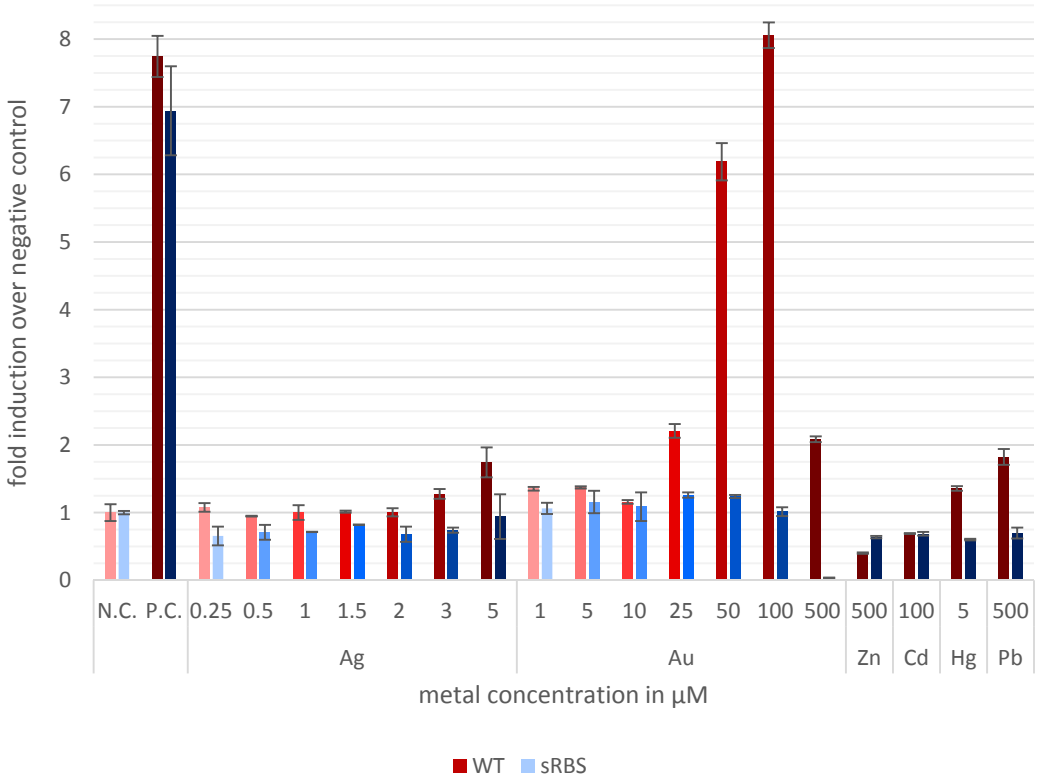

Figure 4.2-17: Metal induction of P_{copA} *in vivo*.
Induction of *sce* expression is shown as fold induction of the metal free negative control. P_{lac} control (P.C.) is shown as benchmark for maximum activity and background activity of the uninduced control (N.C.). Error bars show one SD above and below the average of triplicates.

Compared to the results obtained with the copper promoter P_{copA}, P_{zntA} showed little responsiveness to the tested metals. Zn concentrations above 100 µM reduced fluorescence. Since a repression of the promoter itself is unlikely, these results were interpreted as the result of the strong toxicity of Zn previously observed in the positive control experiments. Lower Zn concentrations perhaps had a weak inducing influence peaking at 1.2-fold for 50 µM Zn, albeit with SDs of almost 0.2-fold above and below this average (figure 4.2-16).

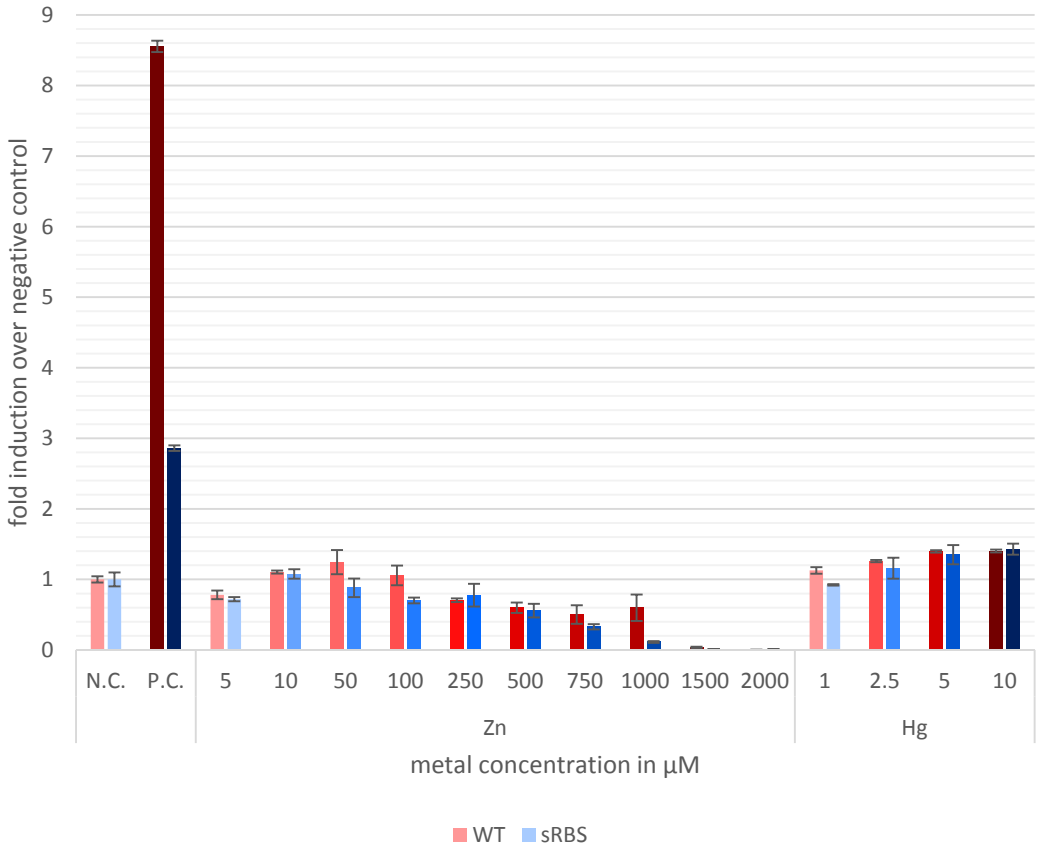

Figure 4.2-18: Zn and Hg induction of P_{zntA} *in vivo*.

Induction of *sce* expression is shown as fold induction of the metal free negative control. P_{lac} control (P.C.) is shown as benchmark for maximum activity and background activity of the uninduced control (N.C.). Error bars show one SD above and below the average of triplicates.

Hg, Cd and Pb had similarly weak effects for the unmodified promoter construct. Background activity of the sRBS version was drastically higher than in the other samples with the positive control only reaching the equivalent of 2.8-fold expression of the uninduced control where this had been 8.6-fold in the native version (figure 4.2-16 and figure 4.2-17). Some 1.6-fold induction through Cd was observed..

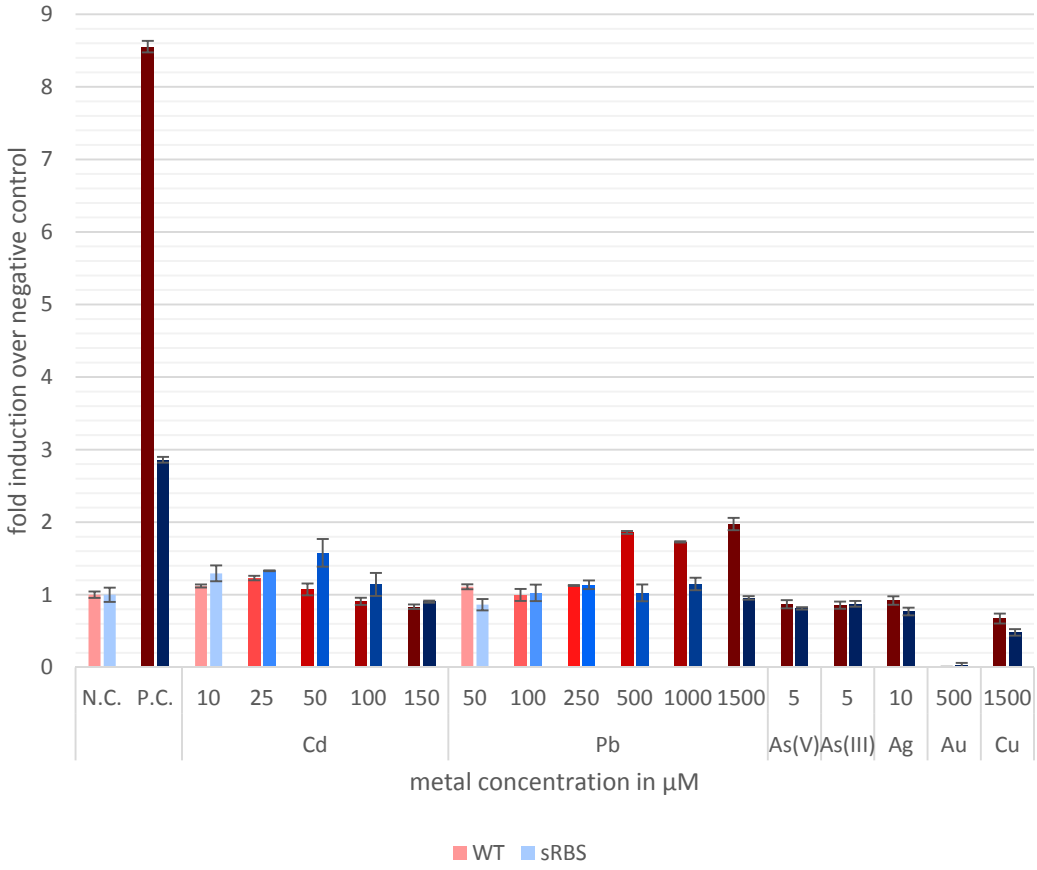

Figure 4.2-19: Metal induction of P_{zntA} *in vivo*.
Induction of *sce* expression is shown as fold induction of the metal free negative control. P_{lac} control (P.C.) is shown as benchmark for maximum activity and background activity of the uninduced control (N.C.). Error bars show one SD above and below the average of triplicates.

Both $P_{merTPAD}$ versions showed no clear tendencies for metal induction by Hg, Cd, Au, As and Ag. Zn (for the native RBS) showed a mild increase in expression up to 1.6-fold at 50 µM before reducing again, somewhat similar to the effects seen in the P_{zntA} experiments. Cu and Pb led to some weak induction at the single concentration tested for them. Background activities were higher for both versions than they had been for the native versions of P_{copA} and P_{zntA} (figure 4.2-18 and figure 4.2-19).

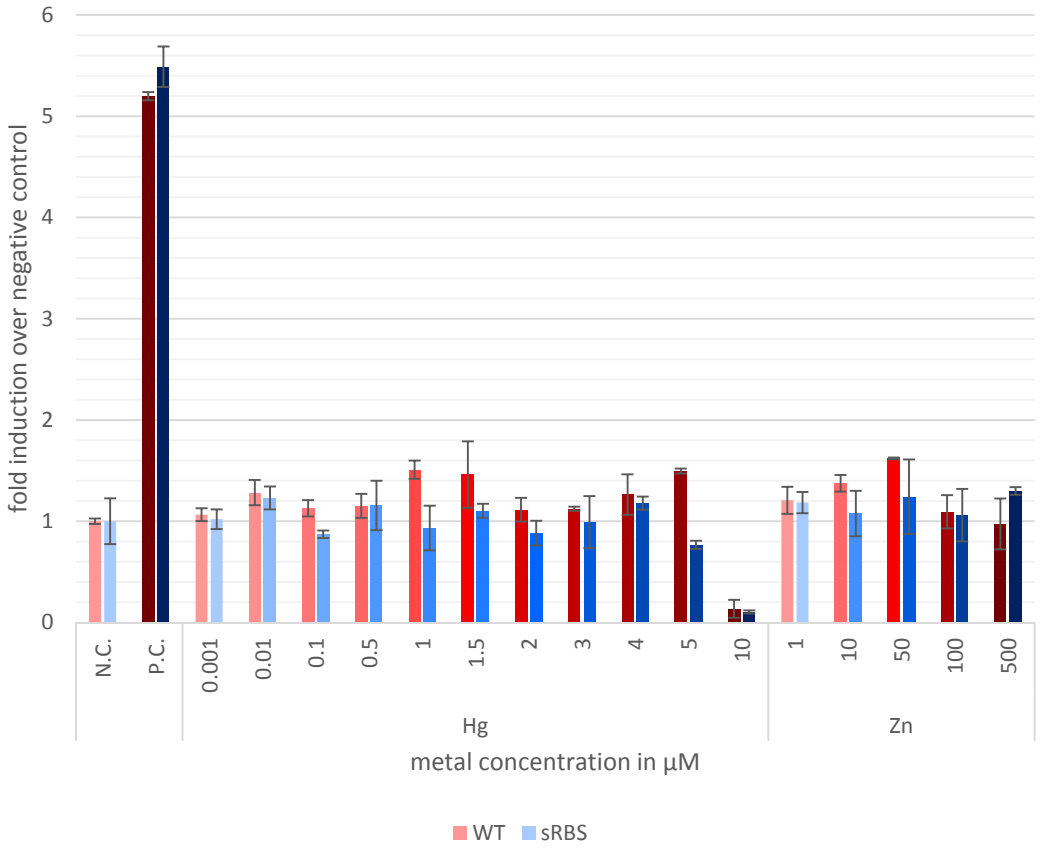

Figure 4.2-20: Hg and Zn induction of $P_{merTPAD}$ *in vivo*.

Induction of *sce* expression is shown as fold induction of the metal free negative control. P_{lac} control (P.C.) is shown as benchmark for maximum activity and background activity of the uninduced control (N.C.). Error bars show one SD above and below the average of triplicates.

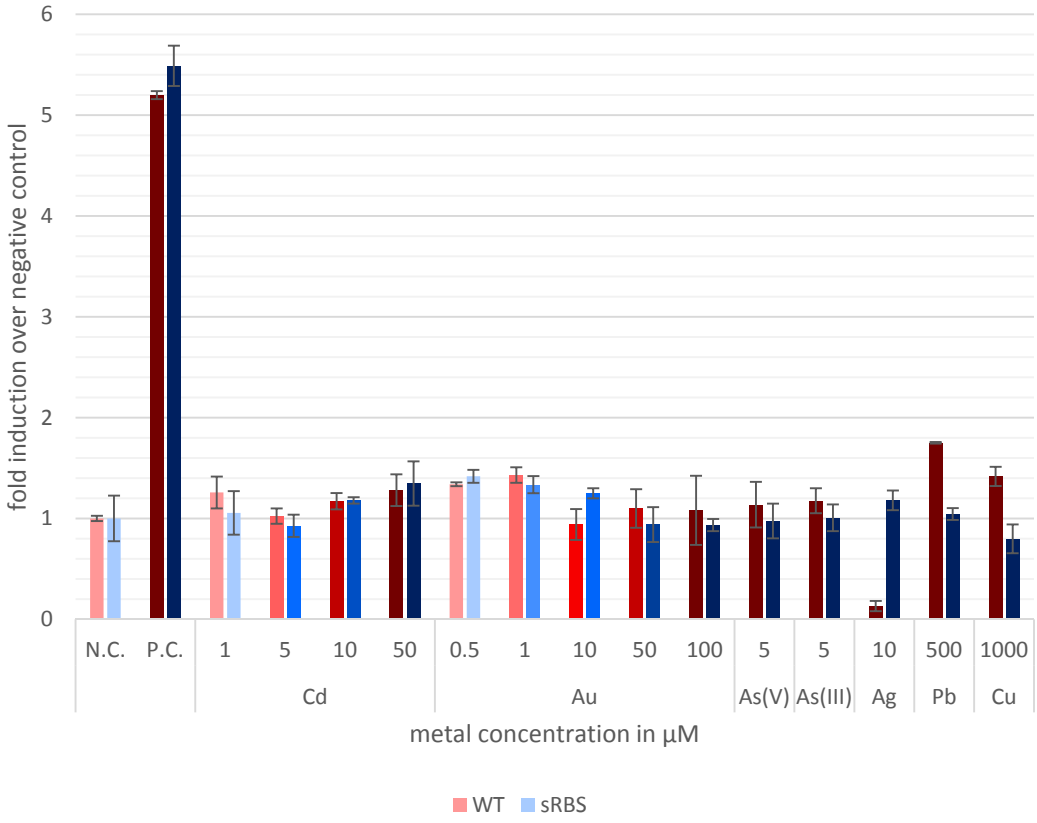

Figure 4.2-21: Metal induction of $P_{merTPAD}$ *in vivo*.

Induction of *sce* expression is shown as fold induction of the metal free negative control. P_{lac} control (P.C.) is shown as benchmark for maximum activity and background activity of the uninduced control (N.C.). Error bars show one SD above and below the average of triplicates.

Due to the different regulatory mechanism of the P_{ars} promoter, testing the metal inductivity *in vivo* without supplying the regulatory protein in sufficient amounts seemed futile. Only the native P_{ars} version was tested *in vivo*. SDs and background activity were generally larger than had been observed for the other promoters, making interpretation problematic. Only the addition of Pb at 1000 µM led to an induction (2.4-fold) standing out from the somewhat fluctuating background induction by the other metals tested. This may be due to the effects of Pb on the cells (figure 4.2-20).

Based on the results from the *in vivo* assays it was decided to not test the $P_{merTPAD}$ construct *in vitro*.

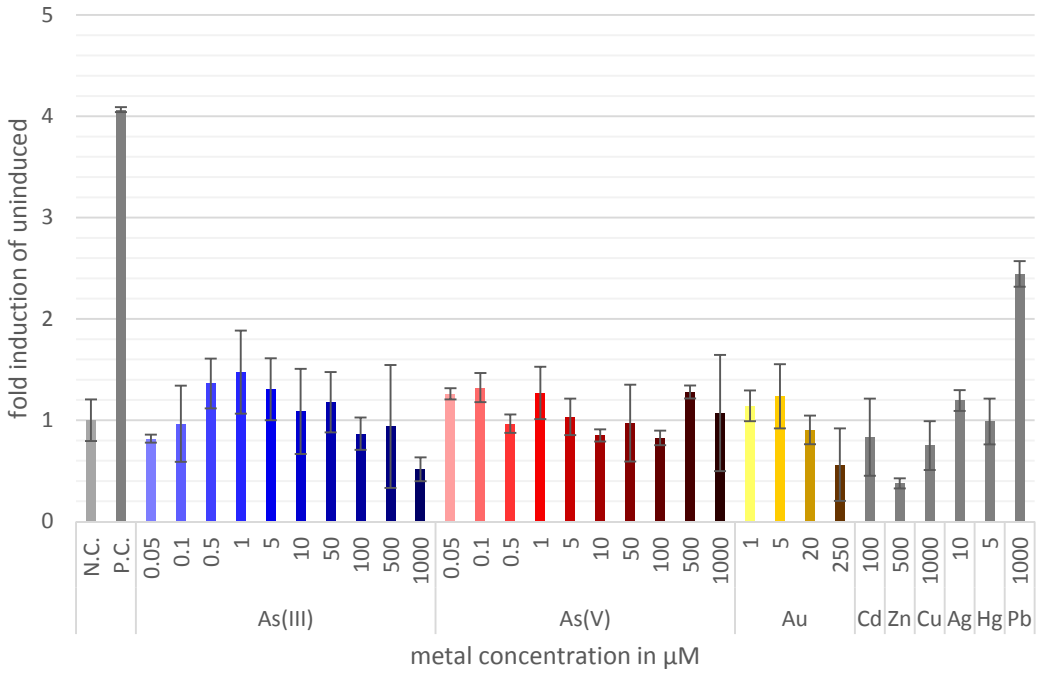

Figure 4.2-22: Metal induction of P_{ars} *in vivo*.

Induction of *sce* expression is shown as fold induction of the metal free negative control. P_{lac} control (P.C.) is shown as benchmark for maximum activity and background activity of the uninduced control (N.C.). Error bars show one SD above and below the average of triplicates.

4.2.5 Supplying the regulatory proteins

In preparation for the *in vitro* testing, the necessary regulatory proteins had to be prepared. The *cueR* and *zntR* genes were cloned into pET28 following a strategy highly similar to the cloning of *arsR* for the AsGard experiments (3.3.1). CueR and ZntR fusions carrying C-terminal His tags were expressed in *E. coli* BL21(DE3) and purified via IMAC (2.2.4.2). ArsR was available as ArsR-His and AGCH from the AsGard experiments.

Samples were taken from the expression cultures for CueR and ZntR at analysed on SDS-PAGE gels. Simultaneously grown cultures with empty vectors and no vector were compared on the same gels to assure that the origin of strong protein bands was the respective fusion protein. Cultures were grown at 30°C and 37°C and as previously observed for ArsR, expression at 30°C seemed more favourable. Figure 4.2-21 and figure 4.2-22 show the SDS-PAGE gels visualising the occurrence of a strong protein band at approximately 46 kDa, the expected protein size, after induction.

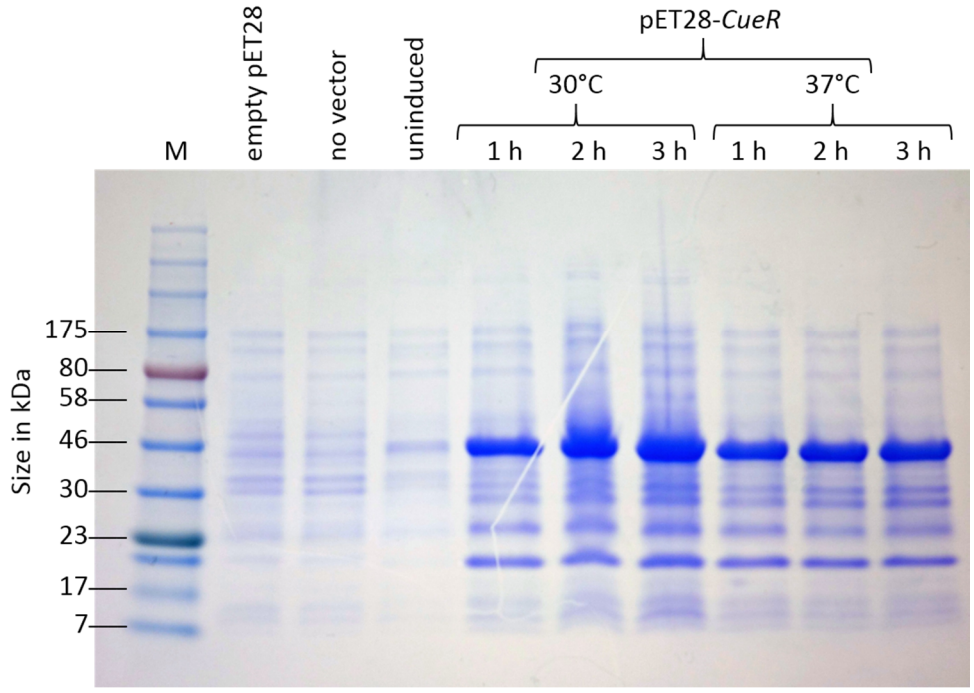

Figure 4.2-23: SDS-PAGE gel of expression culture samples for CueR expression.
M = marker, relevant sizes of the molecular weight marker are indicated on the left.

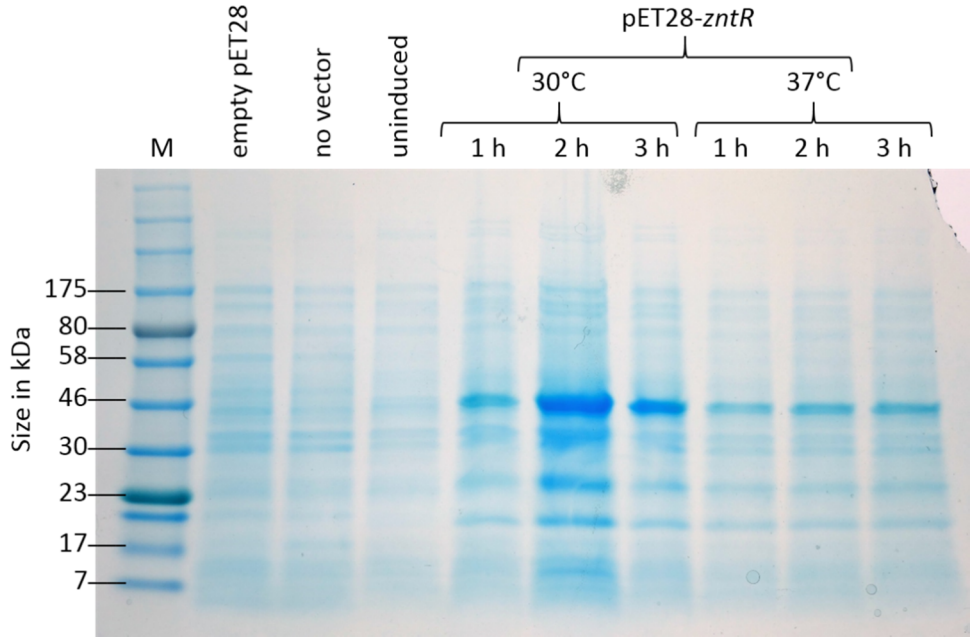

Figure 4.2-24: SDS-PAGE gel of expression culture samples for ZntR expression.
M = marker, relevant sizes of the molecular weight marker are indicated on the left.

Similarly to the samples from the expression cultures and as described previously for ArsR (3.3.1.1), samples from the protein purifications were loaded onto SDS-gels to identify the protein fractions holding the highest concentrations of the desired proteins and losses in the wash fraction. Since the results from the expression culture samples had indicated superior expression at 30°C, only those cultures were used for purification. Both proteins showed some loss in the wash fractions. CueR elution was similar between the three elution fractions. ZntR eluted mainly at 0.1 M and 0.2 M imidazole (figure 4.2-23).

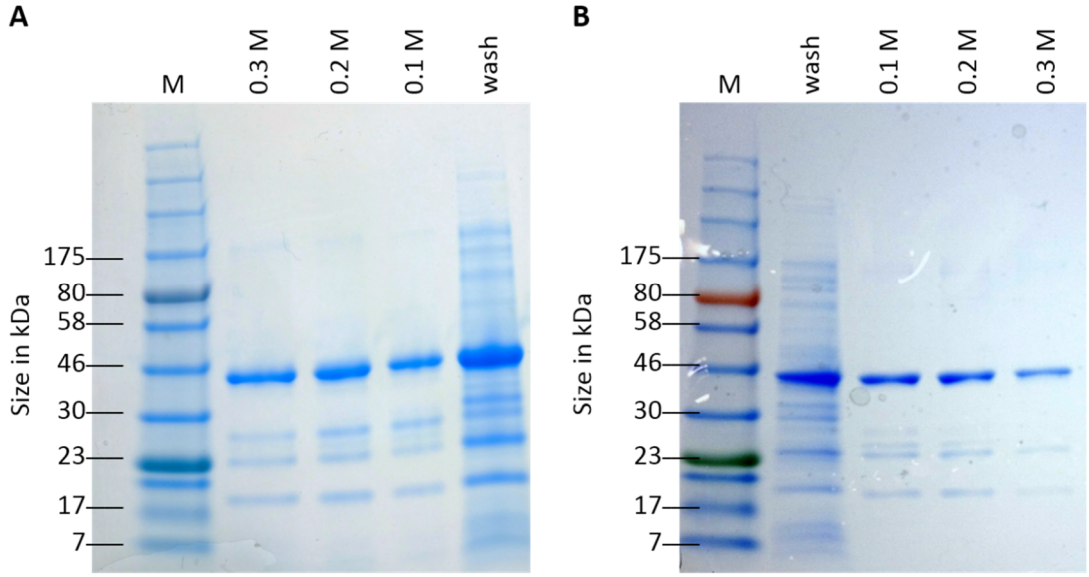

Figure 4.2-25: SDS-PAGE gels of purifications for CueR (A) and ZntR (B).
The wash fraction and three step gradient elution fractions are shown for each protein. Elution took place at 0.1 M, 0.2 M and 0.3 M imidazole. M = marker, relevant sizes of the molecular weight marker are indicated on the left.

To show the unaffected specific DNA-binding capabilities of the proteins, EMSAs were performed as previously described for ArsR (3.3.2 and 3.3.7). For use as specific DNA binding probes, oligonucleotides were designed, covering 38 bp of the binding region of the respective promoter (figure 4.2-24). These probes were named CueR-BS and ZntA-BS respectively. As competitor DNA, the previously described binding probe ArsR-BS was used.

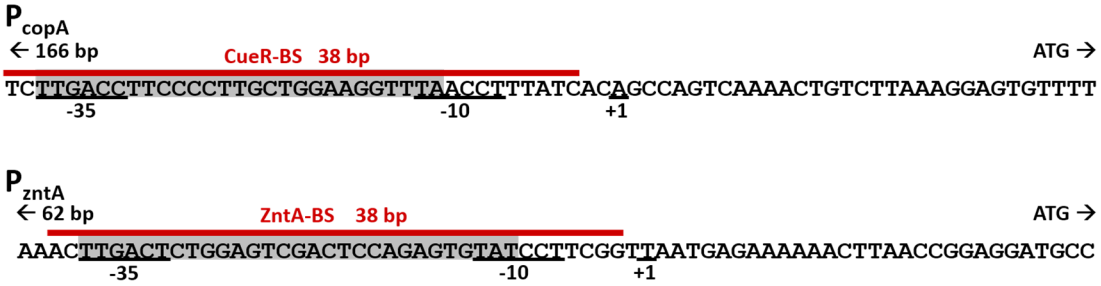

Figure 4.2-26: Coverage of DNA binding probes for CueR and ZntA.

The 38 bp binding probes are shown as a red bar above the partial promoter sequences of P_{copA} and P_{zntA}. The Protein binding region of each promoter is marked by a grey box. Relevant features of the promoters are shown below each sequence.

EMSA experiments revealed unimpeded specific DNA binding activity of both proteins, regardless of the presence of competitor DNA. The addition of heavy metals was not expected to have any effect on the binding activity and results were consistent with this expectation (figure 4.2-25 and figure 4.2-26). No shifting was visible in the presence of competitor probes alone.

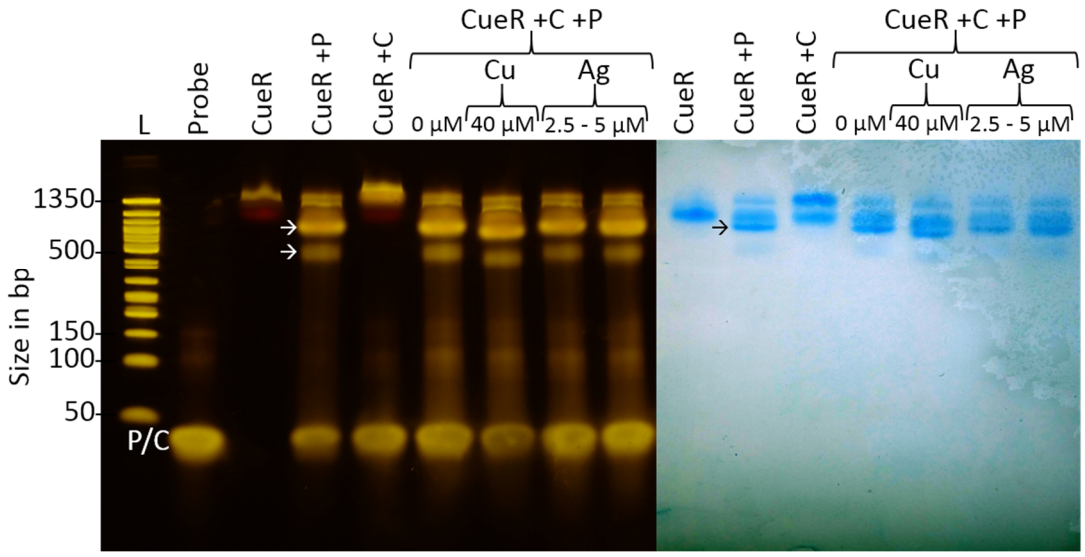

Figure 4.2-27: EMSA of CueR with specific and competing DNA probes.

Two stains of the same precast 5% acrylamide gel are shown, Gel Green staining on the left and Coomassie Brilliant Blue staining on the right. Gel lanes not containing any samples visible with either of the stains have been cut from the images accordingly. Size in bp of relevant bands of the DNA-ladder (L) are indicated. The position of DNA bands of competitor probes (C) and specific DNA probes (P) are labelled. DNA-protein band positions are indicated (→). Approximately 10 fmol of each DNA probe and 100 fmol of protein were mixed with the appropriate metal concentration where applicable and pre incubated in 1x EMSA-loading buffer before loading into the appropriate gel wells.

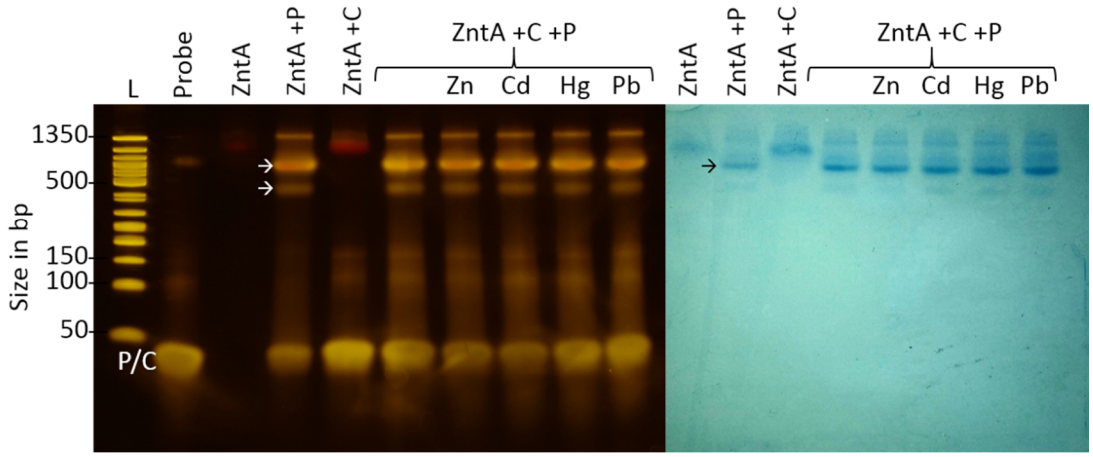

Figure 4.2-28: EMSA of ZntA with specific and competing DNA probes.

Two stains of the same precast 5% acrylamide gel are shown, Gel Green staining on the left and Coomassie Brilliant Blue staining on the right. Gel lanes not containing any samples visible with either of the stains have been cut from the images accordingly. Size in bp of relevant bands of the DNA-ladder (L) are indicated. The position of DNA bands of competitor probes (C) and specific DNA probes (P) are labelled. DNA-protein band positions are indicated (→). Approximately 10 fmol of each DNA probe and 100 fmol of protein were mixed with the appropriate metal concentration where applicable and pre incubated in 1x EMSA-loading buffer before loading into the appropriate gel wells.

4.2.6 Influence of metals on the cell free expression system

Before metal testing could commence in the cell free expression system, the influence of the metals on the system had to be assessed. Similarly to the corresponding *in vivo* experiment, maximum metal levels were based on literature data and on actual sensing needs as defined by the international limits for heavy metals in drinking water (1.1 and 4.3). Due to the limited number of reactions available with the cell free system, metal testing in this experiment had to be limited to a single, maximum concentration for each metal instead of several concentrations each. The same positive control that had been used *in vivo*, P_{lac}-*sce,* was employed. Sce activity was calculated from 4 h experiments similarly to the *in vivo* experiments. No OD-division was performed as this measurement would have required transparent bottom plates and would have been without meaning given there was no bacterial growth. The majority of the tested metals had a limited or no effect

on the assay. Cu and Hg had the smallest influence, Pb, Ag and Cd each led to a slightly higher activity, translating to a 1.2-fold increase when compared to the metal free control. As(III) led to a slight decrease to 0.84-fold, albeit with a somewhat large error when compared to the other measurements. Zn and Au were found to have a drastic influence, reducing expression to approximately half the level of the metal free control (figure 4.2-27).

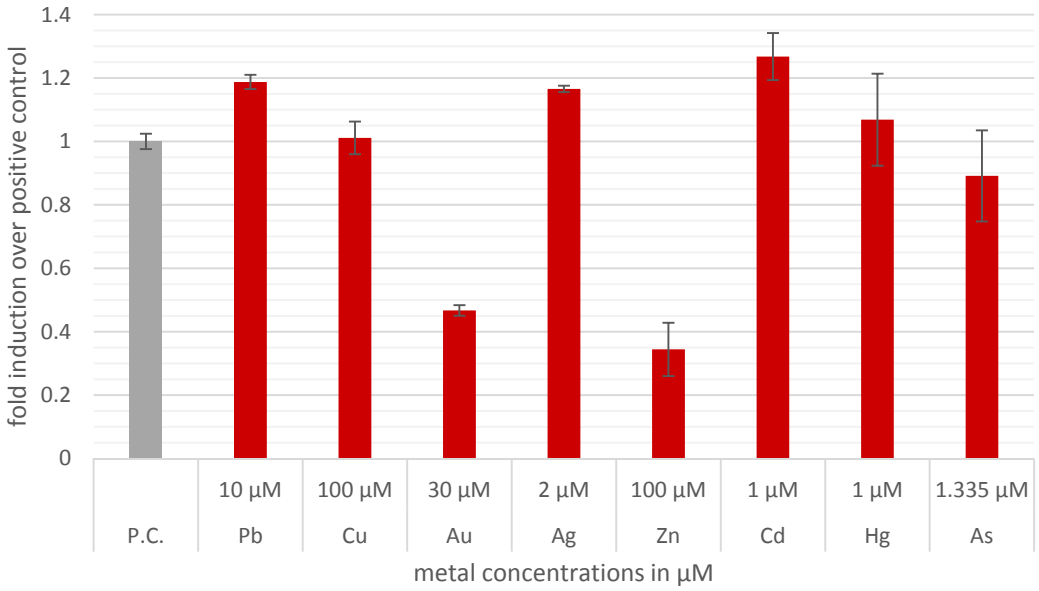

Figure 4.2-29: Influence of heavy metals on the cell free expression system.
Heavy metals were added in the given concentrations to positive control reactions expressing Sce under the control of P_{lac} (data bars shown in red) induced by 0.5 mM IPTG. P.C. denotes the metal free control (data bar shown in grey). Experiments were performed in triplicate, error bars show one SD above and below average.

4.2.7 *In vitro* metal responsiveness

Metal testing was performed for the P_{copA} and P_{zntA}-sRBS promoter constructs. Each reaction was supplied with 6000 fmol of the corresponding regulatory protein (1:50 ratio). Reactions were left to continue for a minimum of 8 h but induction as calculated from fluorescence measurements (figure 4.2-15) was found to be strongest between 2 h and 4 h in most cases.

In case of P_{zntA}-sRBS, Zn, Hg, Cd and Pb were added to the reactions in various concentrations. All inductions were in the range of 0.5-fold to 5-fold that of the metal free control. At the higher end of this range (4.8-fold) was the induction by 10 µM Zn. Increasing the concentration of Zn reduced the observed induction until it vanished completely at 100 µM. For the remaining metals, patterns showed increasing induction with increasing metal concentrations. For all three metals the lowest tested concentrations led to some induction (maximum 1.4-fold), partially with greater errors than had been observed for the other samples. This initial induction remained largely stable until a certain threshold concentration had been reached at which induction started to increase. Concentrations required for Cd and Pb induction were similar, with the first greater increase in induction happening at the step from 0.1 µM to 1 µM metal. For Hg the concentrations were lower, with the increase happening somewhere between 0.025 µM and 0.3 µM.

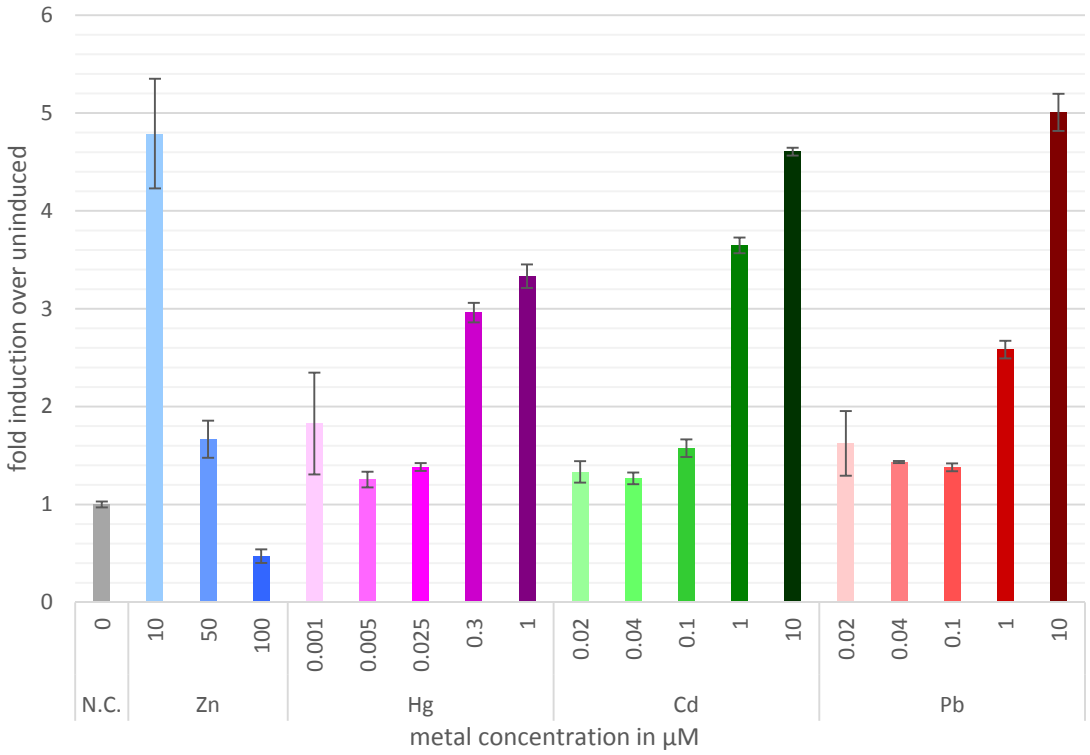

Figure 4.2-30: Induction of P_{zntA}-sRBS *in vitro*.

Heavy metals were added in the given concentrations. P.C. denotes the metal free control (data bar shown in grey). Experiments were performed in triplicate, error bars show one SD above and below average.

Experiments for P_{copA} were performed in a similar manner (except for the different metals employed) but results were discouraging. None of the metals added led to a clear induction or a characteristic pattern. Addition of more than 5 µM Au led to a decrease in fluorescence, corresponding with the results from the metal influence testing. Samples with Ag showed a very weak increase in induction with increasing metal concentrations, yet are far from being more than a vague hint.

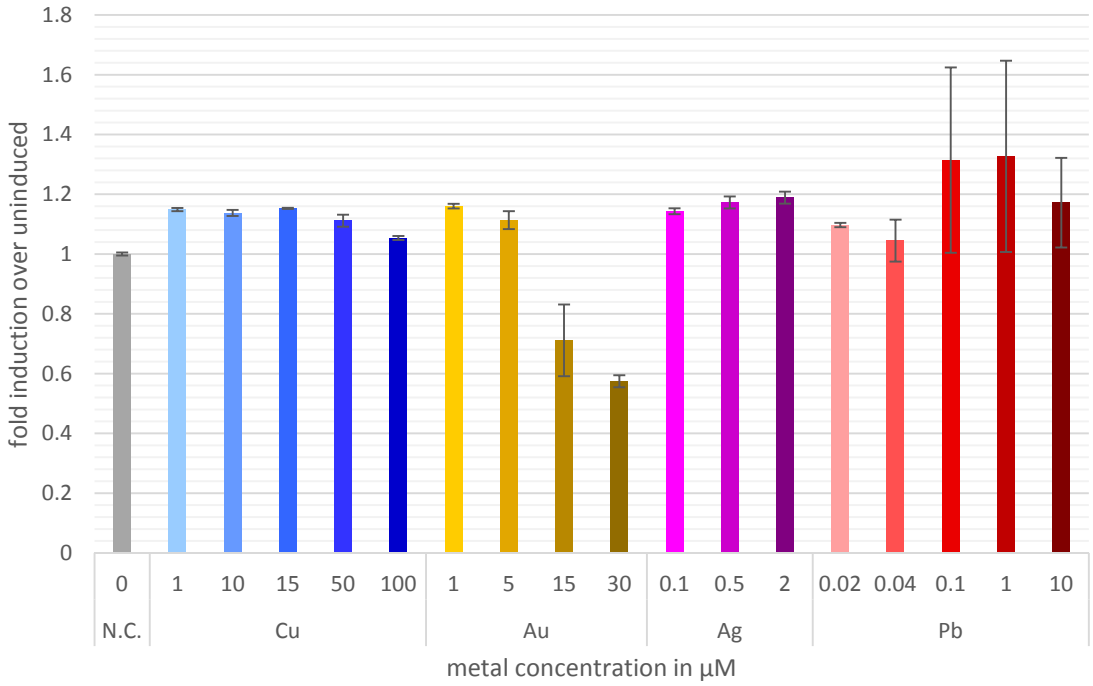

Figure 4.2-31: Induction of P_{copA} *in vitro*.
Heavy metals were added in the given concentrations. P.C. denotes the metal free control (data bar shown in grey). Experiments were performed in triplicate, error bars show one SD above and below average.

4.2.8 A tuneable As(III) response

The ArsR regulation is based on metal unbound ArsR acting exclusively as a repressor while it is attached to the P_{ars} promoter. This mechanism allowed for the design of a tuneable As(III) sensor. Supplying the regulatory protein in a range of ratios to the otherwise unchanged cell free expression reactions containing the according P_{ars}-*sce* plasmid was thought to change the metal responsiveness of the system by making it more or less likely that a metal unbound and thus repressing protein would attach to the P_{ars} DNA. .. The ratios of protein/DNA were chosen to cover a wide range with the highest ratio being 500 proteins per plasmid copy. This ratio was limited by the concentration of the purified protein samples and the available volume in the 10 µL reactions. Attempts to concentrate the protein solutions further led to agglomeration. The tested ratios of DNA:protein (all as molecules/copy number) were 1:0, 1:1, 1:10, 1:100 and 1:500 and the experiments

were carried out independently using the purified ArsR-His protein and the AGCH protein. Each of these ratios was tested with 5 As(III) concentrations, 0 µM, 0.067 µM, 0.134 µM, 0.667 µM and 1.335 µM. This metal range focussed around the WHO recommended limit of 0.134 µM (10 ppb) and the older limit of 0.667 µM (50 ppb). All reactions were performed in triplicate.

For the experiments with ArsR-His, first weak effects became visible at the 1:1 ratio and expression began to increase with the 0.667 µM samples. Increasing the ratio to 1:10 did not change this pattern dramatically except that maximal expression was slightly reduced. The first clear changes occurred in the 1:100 ratio samples. Under these conditions, As(III) concentrations as low as 0.067 µM led to an increase in expression. Overall expression at this ratio was reduced to below 0.8-fold of the samples without regulatory proteins and maximal expression was reached at 0.667 µM. Further increased DNA:protein ratios of 1:500 shifted the pattern slightly. Induction by As(III) began above 0.134 µM for these samples. Maximal expression stayed unchanged and minimal expression was marginally lower than for the 1:100 samples. Over all, the observed pattern concurred with the expectations. Increasing the protein amount while keeping the DNA copy number stable increased the regulatory effect of ArsR on the expression. Higher As(III) concentrations were necessary to relieve this increased ArsR repression effect (figure 4.2-30).

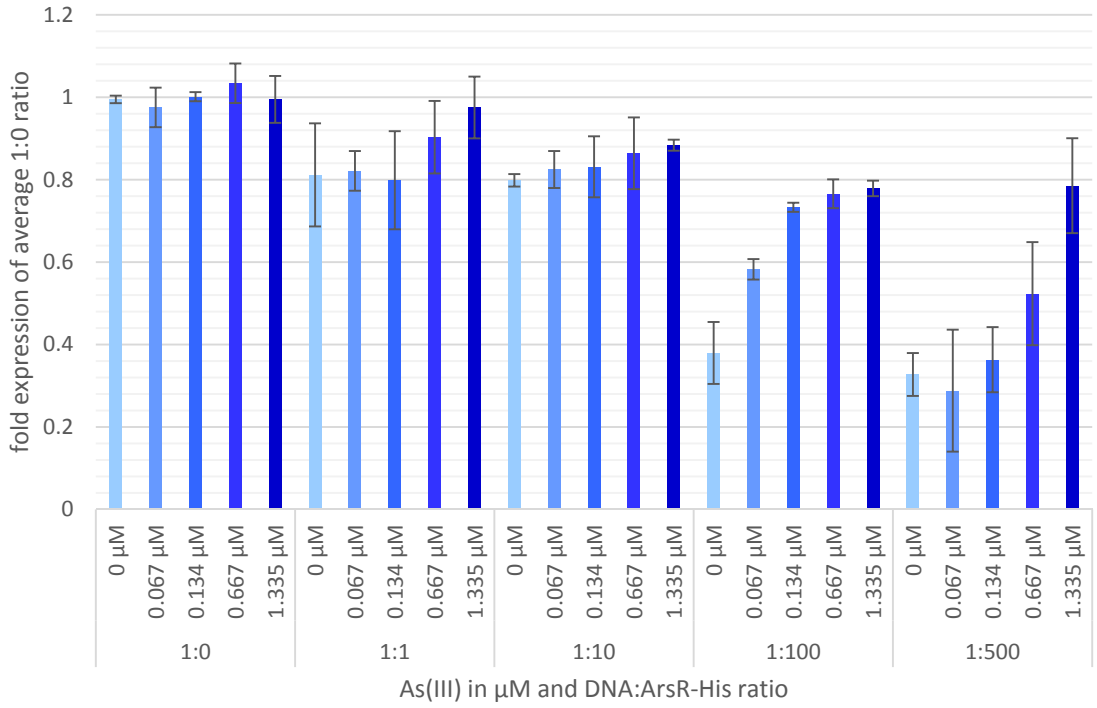

Figure 4.2-32: Responsiveness of P_{ars} in a cell free expression system at different ratios of ArsR-His regulator protein.

Data represents fluorescence at 355/460 nm after negative control subtraction. For each DNA protein ratio, 5 As(III) concentrations were tested. Increasing concentrations of As(III) were coloured with intensifying colours. Error bars represent one SD above and below average of triplicate measurements

This pattern stayed true in experiments performed with AGCH but shifted towards lower protein concentrations, requiring less AGCH than ArsR-His to achieve similar behaviour. Consequently, the highest ratio reduced expression to less than 0.4-fold of the protein free samples and drastically reduced As(III) response even at the highest concentration. Addition of ArsR-His in a ratio between 1:100 and 1:500 led to the best dynamic range while for AGCH the 1:10 ratio seemed more favourable for a similar response pattern.

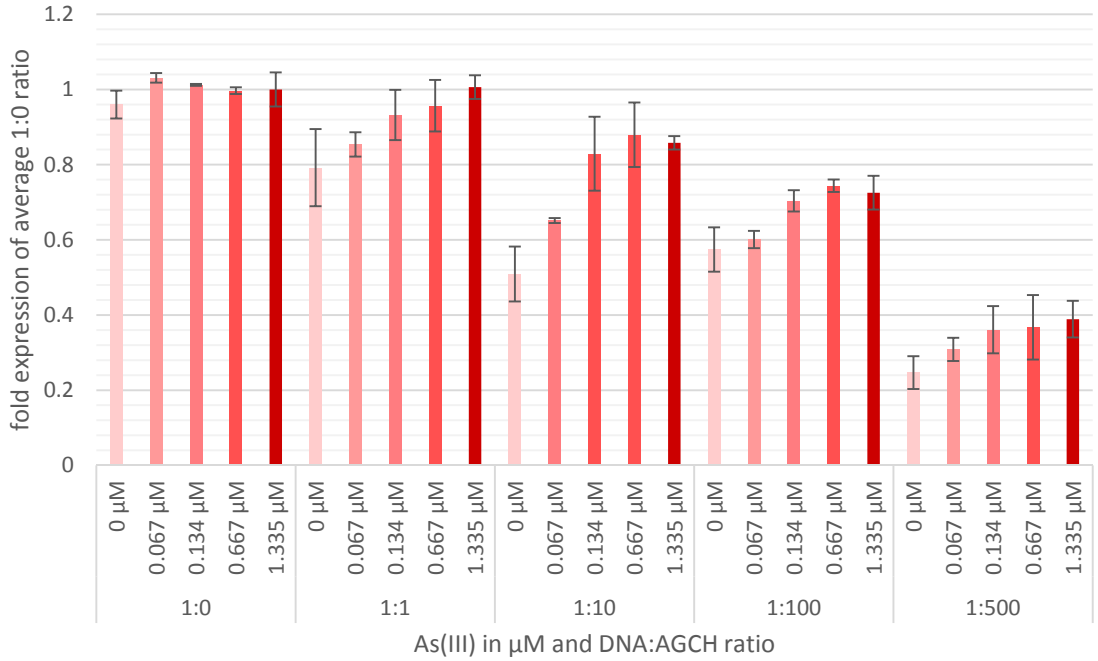

Figure 4.2-33: Responsiveness of P_{ars} in a cell free expression system at different ratios of AGCH regulator protein.

Data represents fluorescence at 355/460 nm after negative control subtraction. For each DNA protein ratio, 5 As(III) concentrations were tested. Increasing concentrations of As(III) were coloured with intensifying colours. Error bars represent one SD above and below average of triplicate measurements.

4.3 Discussion

4.3.1 Reporting, a tough job

To identify an adequate reporter for the metal sensors, several factors were considered. Since the ultimate aim was to bring the sensing system into cell free transcription translation systems, compatibility with these systems was imperative. A pH-change based system would most likely inhibit the expression system massively. A strong signal was desired as the metal responsive promoters have not been reported to possess very high expression levels as are normally favoured in constructs for cell free systems. Lastly the chosen reporter needed to be practical for the high sample numbers to be expected once metal testing would commence. Luciferase is widely accepted as a very sensitive reporter (Su et al. 2011) but it became very clear that the necessary luciferase assay would be impossible to perform with a sufficient reproducibility in the required sample numbers and speed. An advantage of mCherry as a fluorescent protein is the directly proportional feedback. Fluorescence intensity directly reflects the protein copy number and allows for easy derivation of promoter activity over a wide dynamic range.

Unfortunately, the expression levels of the cell free system in conjunction with the metal promoters and even the P_{lac} promoter were too low for this approach. It should be remembered here that with some optimisation and a modified promoter setup, expression levels could perhaps be increased sufficiently for a fluorescent protein approach in the future. However, during the work on this project such optimisation could not be realised and Sce remained a viable combination between the enhanced signal of an enzymatic reporter and the practicality of a reaction that does not require the addition of reagents directly prior to measurement.

Perhaps the biggest drawback of Sce as a reporter is the requirement for fluorescence measurement at UV wavelengths. Qualitatively insufficient plastics as well as autofluorescence from cells, media components and extracts make measurements at these wavelengths prone to high backgrounds or a measurement drift over time as auto fluorescent components are being used up or released.

4.3.2 P_{copA}, P_{zntA} and $P_{merTPAD}$

The *in vivo* experiments had revealed a range of characteristics. Firstly the sRBS version seemed to perform worse in P_{copA}, while with P_{zntA} results were similar between the two versions. Secondly P_{copA} appeared to be the significantly better promoter regarding induction by the tested heavy metals. Keeping in mind that all these assays will have suffered from a severe undersupply of regulatory protein, conclusions should be drawn with care. Both promoters showed a low background activity. As the concentration of the regulatory proteins is unknown, two explanations for the vastly different induction behaviour become apparent. Either CueR and ZntR levels in the cell under the tested conditions are different, leading to different induction levels, or the protein concentrations are similar, entailing a different induction behaviour due to differences in the maximum promoter activity or metal responsiveness. The latter case would explain the low background for both promoters by the presence of repressing regulator, the first scenario relies on a low basal activity of P_{zntA}-WT in absence of ZntR.

With the currently available data neither possibility can be completely excluded. It appears unlikely however that P_{zntA} would show so little induction while ZntR is present in sufficient amounts. Taking the available literature into account, it becomes even more likely that the lack of strong induction is due to a lack of ZntR, as Zn, Cd and Pb would be expected to have strongly inducing effects at the tested concentrations (Brocklehurst et al. 1999; Binet & Poole 2000). Consequently an insufficient level of regulatory protein is the simpler explanation and as such perhaps the one more likely to be true (Occam's razor).

In combination with the data obtained from the P_{copA} experiments, this would mean that the intracellular levels of ZntR are generally lower than the levels of CueR. Only additional experiments analysing the actual RNA and protein levels could elucidate this. In any case, the data obtained with the P_{zntA} indicated some weak induction by Hg and possibly Pb. It should be noted here that both metals had shown a slightly enhancing effect on fluorescence development *in vitro* and that this may be the actual cause for the increased fluorescence.

The P_{copA}-WT experiments behaved largely as expected and showed surprisingly good induction for Cu and Au (given the expected undersupply of CueR). When

comparing the results obtained in this work with data available from literature, it becomes apparent that the sensing construct is working as intended and that increased supply of CueR would likely lower the sensing thresholds and increase the feedback strength at higher metal concentrations. Looking at the literature, Outten et al. relied on the natively present CueR to control a single plasmid bound copy of P_{copA}, while Stoyanov et al. employed a two plasmid system with a higher copy number for the CueR expressing plasmid (Outten et al. 2000; Stoyanov et al. 2001). The lower Cu concentration necessary for maximum induction in the experiments carried out by Stoyanov et al. may reflect this difference in experimental design and emphasises the importance of sufficient CueR levels. Interestingly, the experiments in this work allowed for much higher Cu concentrations without toxicity than any of the aforementioned publications reported. It is possible that this difference lies in the use of LB instead of minimal medium. LB may contain some chelating agents and may give better support for cells struggling to cope with heavy metal influences. Similar findings were reported by Stoyanov et al., alas no data was published (Stoyanov et al. 2003). Interestingly, the very strong induction for Au observed in this work had previously been reported to be more linear and overall weaker (Stoyanov et al. 2003, table 4.3-1). Summarising, it can be said that the Cu sensing construct performed as well as was to be expected without the supply of additional CueR.

Table 4.3-1: Comparison of *in vivo* activation thresholds of P_{copA} for selected metals between literature data and this work.

All concentrations are given in µM. Where available the first metal concentration showing a difference to basal activity and the maximum induction concentration are given. Where no induction was observed fields are marked "none". Where no such experiment was carried out, fields are marked with "NA". Values marked with an asterisk indicate possible hypersensitive response curves. Values marked with a + indicate that the maximum activation concentration may be higher than the highest tested concentration (Outten et al. 2000; Stoyanov et al. 2001; Stoyanov et al. 2003).

Source	Point	Cu	Ag	Au	Hg	Pb	Cd	Zn
Outten 2000	start	10	NA	NA	none	NA	NA	none
	maximum	500						
Stoyanov 2001	start	1	1.25*	NA	none	none	none	none
	maximum	50	2.5*					
Stoyanov 2003	start	1	0.1*	20	NA	NA	NA	NA
	maximum	60	0.75*	40				
This work	start	1	3*	1*	5	500	none	none
	maximum	1500 +	5+*	100*	NA	NA		

The $P_{merTPAD}$ promoter did not show any induction and gave a higher background, both with WT RBS and sRBS,when compared to the other two promoters. Due to the limited sample number available as cell free reactions, it was decided to postpone work on $P_{merTPAD}$ until a self-manufactured cell free expression system of sufficient fidelity would be available.

Moving on to the *in vitro* experiments, the P_{copA} – P_{zntA} situation inverted. P_{zntA}-sRBS showed clear induction by all metals tested, indicating that the system of supplying the regulatory protein as an ingredient of the reaction rather than expressing it within, is a viable strategy. Furthermore, it shows that the P_{zntA}-sRBS construct can function as intended and that the *in vivo* results for the engineered version are not influenced by a dysfunctional sensor construct. As expected, Zn led to some induction but toxic effects make the sensing of relevant concentrations around and above 50 µM impractical. Since Zn is only considered in drinking water for cosmetic reasons, Zn

sensing is of low priority. More importantly, Hg, Cd and Pb were successfully sensed, not much above their legal or recommended limits in drinking water (table 4.3-2). Minor improvements to the assay could possibly lead to the sensing thresholds dropping into the relevant range.

Table 4.3-2: Comparison of heavy metal limits and sensing thresholds of the cell free sensing system with P_{zntA}-sRBS.
All values are given in µM.

Metal [µM]	**Cd**	**Hg**	**Pb**	**Zn**
Limit range	0.018 - 0.046	0.005 - 0.010	0.05 - 0.07	45.9 - 76.5
Sensing threshold	>0.4	>0.025	>0.1	<50

The first approach that springs to mind is the variation of the supplied ZntR amount. Promoter copies that are not bound to ZntR are not available for the sensing process, reducing output strength. However, over-supply could at some point lead to the high concentrations of unbound protein competing for metal ligands with the DNA bound protein copies, thus reducing the available metal concentration for the actively sensing complexes. With the current data it is not possible to anticipate if this level may have been reached with the 50-fold supply. Apart from this fine tuning of the sensing mechanism, an increase in expression once induced might also facilitate lower sensing thresholds. A promising approach might be the addition of a T7 promoter in front of the current sensor construct, a strategy that was successfully employed by the 2015 Bielefeld iGem team (Bielefeld CeBiTec). This addition could increase the overall expression levels, albeit an increased background activity might counteract this advantage. Alternatively and somewhat more adventurous, is an approach based on a report by Brown et al. in 2003. They were able to show that domain swaps between MerR-like regulators can produce regulatory elements with different promoter recognition properties (Brown et al. 2003). While their conclusion was simply that this shows the similarities within the family, a biosensing point of view makes this finding a first step in the construction of chimeric proteins. Employing the best DNA binding and regulation capabilities of one protein and combining them with metal sensing domains of interest, could give rise to super

proteins combining the advantages of multiple proteins into one for maximum sensitivity. Creating sets of such chimeric proteins would allow for a vastly extended sensing range when compared to the natively available proteins.

It remains unclear why the P_{cop}-WT construct showed no induction in the cell free system. Since it had functioned comparatively well *in vivo*, it is not likely that the construct itself is the cause of this problem. A possible reason would be insufficient quality of the supplied CueR. While the regulatory proteins had been tested for their DNA binding properties via EMSA, this does not necessarily indicate their consistent functioning as regulatory elements. In addition EMSA had revealed some influence of added competitor DNA, although not in the form of clear shifting. Change of the CueR expression and purification routine might be necessary. The aforementioned works of Outten et al. and Stoyanov et al. suggest that a limiting factor in sensing with these promoters is the regulator (Outten et al. 2000; Stoyanov et al. 2001). Consequently, besides the quality of the supplied CueR proteins, the required amount for effective regulation and sensing might differ from ZntR and finding the right ratio requires further testing.

Due to the inconsistent behaviour of the sRBS versions in comparison with the promoter with unchanged RBS, drawing final conclusions regarding advantages and disadvantages of these changes becomes speculative. They may indicate an insufficient precision in the predicting algorithm employed or indicate that previously undescribed promoter elements or protein interaction domains may be affected by the sequence change. Data available for P_{ars} (figure 3.4-1, Xu et al. 1996; Siddiki et al. 2011) indicates that the identification of protein-DNA binding sites and contact points in the available publications should be regarded with care, as different publications proclaim different localisations of these sites. However, the differing behaviour between the *in vivo* and *in vitro* experiments and the limited sample number make the formation of a clear hypothesis unrealistic. Further testing would be required to resolve the actual effects of the undoubtedly daring changes.

4.3.3 Tuneable sensors

The concept of a tuneable As sensor has first been realised *in vivo* to some extend by Merulla et al. in 2013. The i*n vitro* system presented in this work exhibited similar shifts in sensitivity upon change of the regulator-DNA ratio. Inclusion of AGCH showed that the increased DNA binding strength observed in the AsGard project indeed led to a tighter regulation and a lower As(III) sensitivity of the sensor when compared to ArsR-His. Combining the effects of the DNA-regulator ratio and the fusion-protein characteristics, presents an excellent toolset for the creation of fully tuneable As sensors that can be very sensitive on one hand and cover a wide range of concentrations on the other hand. In contrast to the work published by Merulla et al., the cell free system with the separate supply of regulator and DNA allows for a precise adjustment of ratios. Merulla et al. did not publish any RNA or transcriptome analysis and the promoters used for the expression of ArsR were only identified by their relative activity to each other, not their absolute activity (Merulla et al. 2013). Hence this work presents the first data for such a tuneable system for As where absolute ratios become available that are reproducible and can be standardised. With a deepening understanding of the specifity of metal recognition by metalloregulatory proteins, it may become possible to apply this strategy to other regulators by combining their metal recognition sites with the regulatory sites of SmtB-ArsR family proteins with the desired affinities for their respective DNA binding partners. Already the presented sensing system is able to detect As(III) in the relevant range of 0.134 µM to 0.667 µM (equivalent 10 ppb – 50 ppb). Testing with a higher resolution of metal concentrations would allow for the calculation of the relevant dissociation constants and standard curves for metal response vs DNA-regulator ratio. Combination of the regulatory region of the P_{ars} promoter with a T7 promoter might, as with P_{copA} and P_{zntA}, provide an enhanced signal and an overall better performing sensor.

4.4 Conclusion

Striving for a multi-metal sensing all-in-one biosensor, a set of novel cell free biosensors for heavy metals has been presented. In their modular design for use in cell free expression systems, these sensors aim to address the limitations of whole-cell biosensing and take advantage of the defined, standardised characteristics of cell free expression systems. A P_{zntA} based sensor has been shown to detect the heavy metals cadmium, mercury, lead and zinc close to their recommended limits in drinking water and data suggests that simple optimisations could further improve this sensitivity. For the first time a precisely tuneable arsenic sensor has been presented. Based on P_{ars}, this sensor can be adapted for individual sensing needs by varying the concentration of the regulator ArsR. This sensor covers the relevant range of arsenic as defined by WHO drinking water guidelines. A selection of optimisations has been suggested to enhance the performance of the aforementioned sensors and two additional sensors based on P_{copA} and $P_{merTPAD}$.

5 Production of cell free expression systems

5.1 Introduction and aims of work presented in this chapter

Cell free expression systems were a crucial tool in this work. While commercial systems are widely available and are quick and efficient, their cost per reaction usually lies around £ 12 (Promega Corporation 2015a). Reducing the sample volume can lower this cost to around £ 2.50 but performing large scale experiments covering a wider range of conditions, as is required in biosensor testing, remains a financial burden. Consequently the aims of this chapter were:

- To establish the production of in house cell extracts for cell free protein expression.
- To improve extract performance to allow larger scale testing of the biosensors created.

5.2 Results

The in house preparation of cell free transcription translation systems was performed according to the most recently published method at the time by Kwon and Jewett (2015). This method was thought to represent the latest advances in the preparation of cell-free extracts and is significantly less time consuming than previous methods (figure 5.2-1). Like all major preparation methods it is based on the careful lysis of *E. coli* BL21 cells suspended in a Tris buffer and subsequent centrifugation to separate cell debris detrimental for activity of the extracts. The buffers and supplements used in the later expression procedures are very similar between the various methods published (Bernhard & Tozawa 2013; Kwon & Jewett 2015). Details of the buffers employed are given in section 2.2.7.

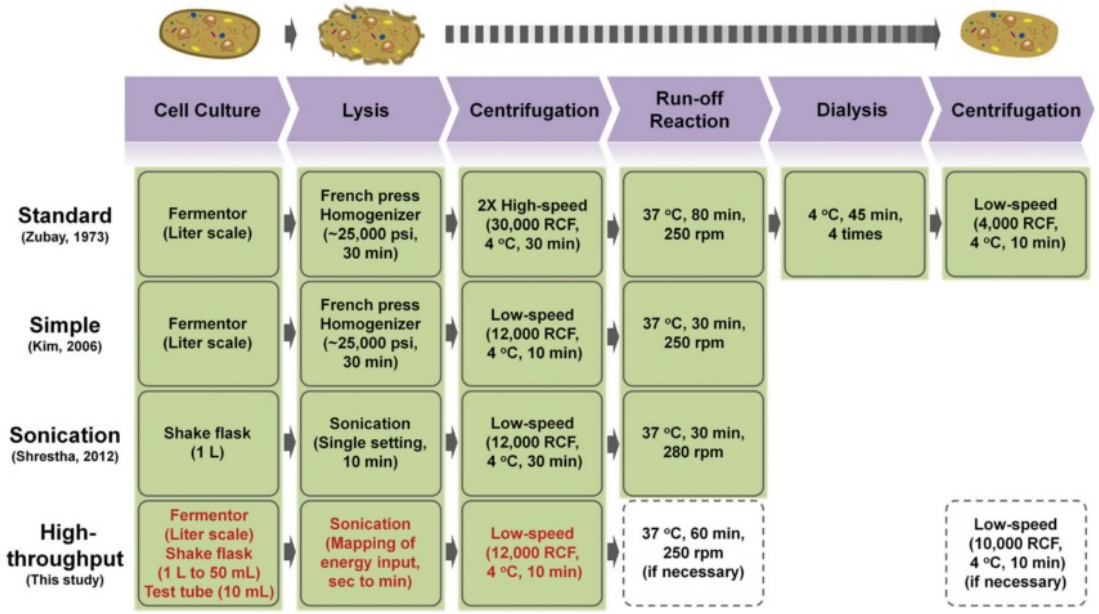

Figure 5.2-1: Comparison of commonly used protocols for cell free expression extract preparation (From Kwon & Jewett 2015).

The power output of the sonicator employed was measured and calculated calorimetrically according to the recommendations by the US National Institute of Standards and Technology (Hackley & Wiesner 2010). A 600 mL cylindrical borosilicate beaker was filled with 500 mL de-ionised water (by weight). The beaker was covered with plastic film and left in the laboratory next to the sonicator for 4 h to allow the temperature of the beaker to equalise with the room temperature. The

plastic film was then removed and the air temperature of the room and the water were compared. The experiment was not continued if the temperature was not found to be identical. The sonicator probe was immersed approximately 2.5 cm into the liquid and a digital temperature probe was immersed approximately 1 cm away from the sonicator probe. Using the sonicator setting later used in sample sonication, temperature changes were logged over a 5 minute continuous sonication. The procedure was repeated three times, using a fresh beaker and room temperature water each time. The resulting measurements were used to draw a temperature vs. time curve and a best linear fit to this curve was calculated using the least squares regression fit. With the slope obtained the delivered power was calculated following the formula: $P = \frac{dT}{dt} MC_P$, where P is the delivered acoustic power (W), T is the temperature (K), t is the time (s), M is the mass of liquid (g) and C_p is the specific heat of the liquid (J/g×K) (Hackley & Wiesner 2010). This method bears the usual shortfalls of calorimetric methods in terms of heat loss to the surrounding air and the sample holder, potentially leading to marginally lower than real results for the delivered power. Since the overall temperature increase during the 5 minute period was generally below 5 K, the effects of heat loss will be minimal. It should be noted that the sonicator probe was close to reaching the end of its lifetime and that repeated measurements over the course of several months indicated a deterioration of power output at otherwise unchanged settings from 22.37 W/s to 17.98 W/s.

A total of 6 batches of cell extract were prepared, addressing experiences and complications from each preparation in the following batch. The initial preparation gave rise to extracts with a protein content of approx. 30 mg/mL, below the recommended 40 mg/mL and was prepared using a mixture of amino acids (AAs) prepared in the laboratory (2.2.7, Kwon & Jewett 2015). Activity was assessed with the luciferase positive control (pBest-*luc*) from the Promega S30 circular DNA cell free expression kit after 4 h incubation (figure 5.2-3, data point A). Analysis of the reaction by SDS-PAGE revealed that a considerable amount of a protein consistent with the size of luciferase (61 kDa) had been expressed. This protein was not present in controls performed without the pBest-*luc* plasmid (figure 5.2-2).

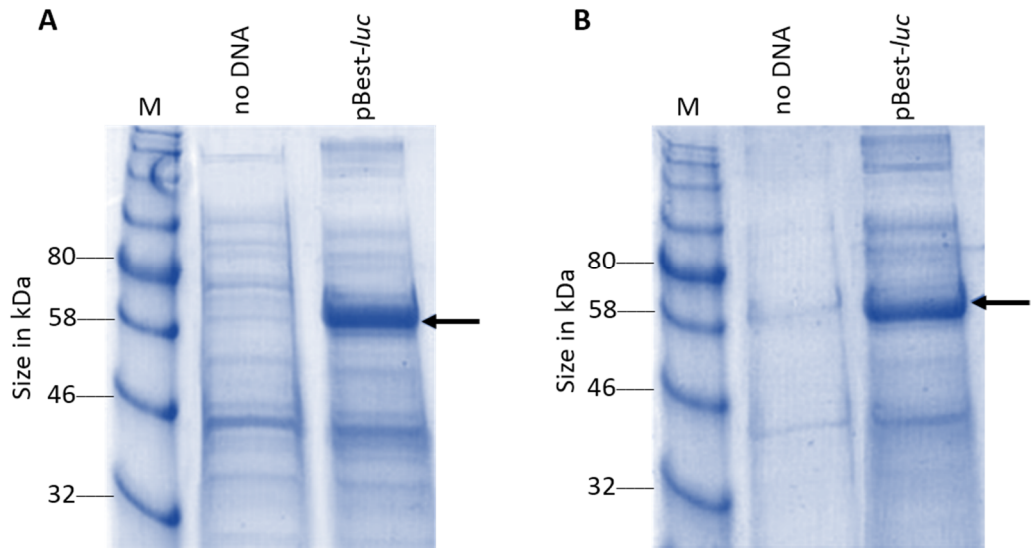

Figure 5.2-2: SDS-PAGE gels of cell free reactions.
For the commercial extract obtained from Promega (A) and the first in house extract (B) samples are shown for a 4 h reaction with pBest-*luc* positive control plasmid and without added DNA. Relevant weights of the protein standard are given on the left of each picture. Gels were stained with Coomassie Brilliant Blue.

Subsequently, the AA mixture from the commercial kit was used instead of the mixture prepared in-house and all remaining buffers were prepared afresh to exclude inadequacy of these mixtures as a possible cause for the very low activities observed. Later batches showed greatly increased protein concentrations, mainly due to improved handling techniques prior to sonication. According to the protocol employed, the cell pellets obtained after centrifugation of the liquid cultures had to be resuspended in very low volumes of buffer prior to sonication. In the first preparation this led to a considerable loss of cell mass due to adhesion to tube walls, pipette tips and other items used during resuspension and liquid handling. Once some experience had been gained as to how to minimise losses during these necessary steps, protein concentrations increased. In an attempt to identify the limiting component in the comparatively low activity of the self-prepared extracts, batch C was tested with the reaction buffers from Promega. This change was found to have little effect. As protein concentrations were higher than recommended, the assumption was made that this excessive protein content might have inhibiting effects on luciferase activity or expression (figure 5.2-3, data points B to E).

Accordingly, the buffer volume for resuspension of the pelleted bacterial culture was increased from 1 mL per g wet cell mass to 2 mL/g, despite this increase not reflecting the recommendations by Kwon & Jewett (Kwon & Jewett 2015). The results indicate that this adaption had a beneficial effect on the extract activity after 1 h incubation but not after 4 h (figure 5.2-3, data point F). Table 5.2-1 gives an overview of the AA mix, reaction buffer and resuspension volume used in each batch. The highest activities achieved with self-prepared extracts in batch F reached approximately a quarter of the activity of the commercially available extracts from Promega.

Table 5.2-1: Overview of buffer, AA mix and resuspension volume used in each batch.
For the AA mix and the reaction buffer "own" is written for each batch where where the self-made solutions were used and "Promega" whenever buffers from the commercial kit were used instead. The resuspension volume is given for each batch in mL/g wet cell weight, the recommendation by Kwon & Jewett (2015) is 1 mL/g. Additional comments are given where relevant.

Batch	**A**	**B**	**C**	**D**	**E**	**F**
AA mix	own	Promega	Promega	Promega	Promega	Promega
Reaction Buffer	own	own	Promega	Promega	own	own
Resuspension volume mL/g	1	1	1	1	1	2
Comments	High cell mass loss	Prepared together, reaction buffer prepared freshly		Prepared together, reaction buffer frozen (-80°C) from previous batches		Reaction buffer prepared freshly

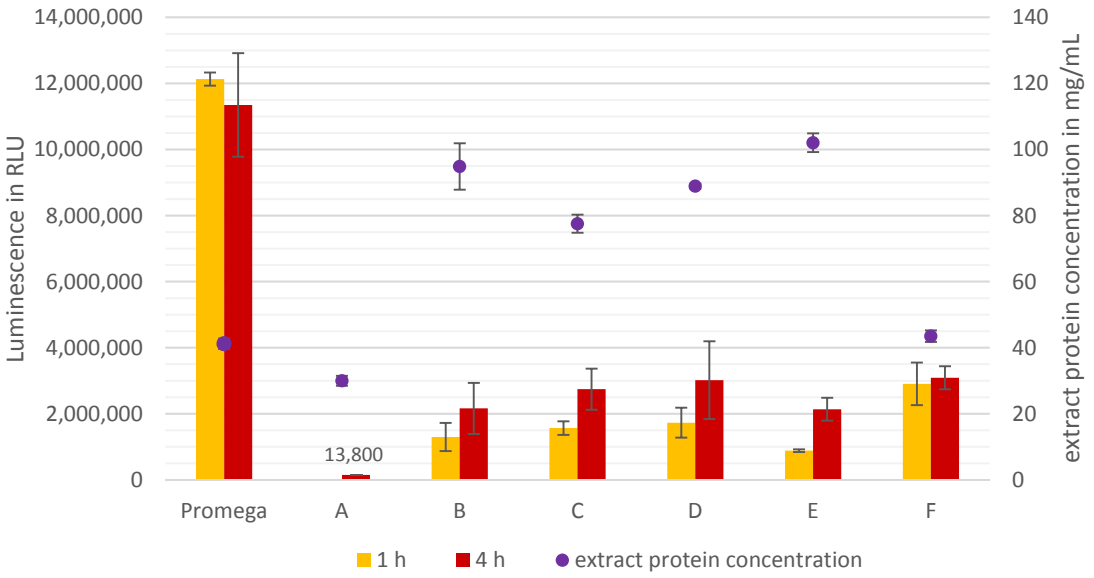

Figure 5.2-3: Luminescence and protein concentration of in house prepared cell free expression systems in comparison to commercial extracts.

Data is shown for 6 extract batches labelled as A to F. Luminescence for batch A was only measured after 4 h and results are shown as numerical value due to limited readability of data bar. Data shown reflects the average of at least triplicate experiments and error bars show one SD above and below average. Negative controls without added DNA showed luminescence values below 100 RLU at all times.

5.3 Discussion

While being significantly easier to perform than earlier methods published, the Kwon & Jewett method of cell free expression extract preparation proved to be problematic. The activities achieved rose with every preparation (except preparation E) but remained unfeasibly low for use of the extracts with the sensing plasmids described previously. Interestingly the SDS-PAGE results for the first batch indicate that large amounts of protein, presumably luciferase, are being produced by the extracts. As such, the lack of luminescence in this preparation seems to be not due to absence of the enzyme but rather its misfolding or otherwise impeded activity. Misfolding could, among other reasons, be the consequence of insufficient quality of the buffers. Unsuitability of the AA mixture may have a detrimental effect on the protein quality and overall activity of the extract (Sun et al. 2013). However, since the handling had improved between batches A and B and the remaining buffers were prepared freshly, the AA mixture is not the only possible cause for the low activity in batch A. The reaction buffers used in batch B and the following did not seem to limit the activity as their replacement with the Promega buffers did not lead to drastically changed activities. Storage of the buffers may have a detrimental effect, as batch E, tested with buffers that had been kept at -80°C, did show a slightly reduced activity.

The method employed does not usually recommend the introduction of a run-off reaction in the preparation process of T7 polymerase based extracts (Kwon & Jewett 2015). In a run-off reaction, the freshly prepared extracts are incubated at 37°C under agitation for 30 to 80 min depending on the protocol employed. After this incubation the extracts are centrifuged at 10,000 x g or above for 10 min or longer at 4°C (Zubay 1973; Kigawa et al. 2004; Shin & Noireaux 2010; Kwon & Jewett 2015). It has been reported that the centrifugation step after this incubation is crucial to extract activity (Kwon & Jewett 2015). This indicates that during the incubation, components of the extract agglomerate or form complexes, which are subsequently removed by the centrifugation.

The data obtained during the aforementioned experiments suggests that such a reaction might be beneficial. When comparing the activity of the commercial extracts with the in house extracts, it becomes apparent that the commercial kit does not profit from incubation times as long as 4 h, while all self-prepared batches with the

exception of batch F showed higher activities after 4 h. The results presented indicate that the transcription-translation capacity of the extracts is not increasing past a certain protein concentration but instead is slowed down due to unknown inhibitory processes. This becomes most apparent when comparing batches D and F. Reducing the protein concentration seems to leave sufficient active components for the transcription-translation apparatus to achieve similar maximum activity levels but reduces the restrictive effects, thus resulting in more of the observable expression occurring within the first hour in batch F than in batch D. It is not unlikely that a part of the inhibiting or interfering components could be removed from the extracts during the run-off reaction and that the remaining extract would exhibit an increased reaction speed and perhaps a greater activity as well. This is based on the assumption that the limiting factor in the self-made extracts is not depletion of certain components but their hindrance or loss of activity. Kwon & Jewett concluded in their work that a run-off reaction had no beneficial effects on extracts prepared from *E. coli* BL21 Star cells, but their conclusion is entirely based on data collected after 4 h of incubation and with systems employing T7 RNA polymerases (Kwon & Jewett 2015). In any case, the results of the SDS-PAGE for batch A show that it may be the functionality of the luciferase and not the transcription-translation itself that limits the cell free expression system. If SDS-PAGE experiments were to show the same results for the later batches, future optimisation should focus on the ideal conditions for correct protein folding, maturation and activity.

The greatest factor of uncertainty in the preparation remains the sonication. Despite the precise calculation of sonication times based on wet cell weight as described by Kwon & Jewett, inconsistencies in sonicator output and the general variability of sonication seem to make this method of cell disruption prone to precision and reproducibility deficiencies (Kwon & Jewett 2015). Power output, sample volume, sample density, placement of the probe in the vessel, vessel shape, vessel material, medium surrounding the vessel and probe geometry as well as sonicator amplitude can have significant effects on the sonication itself (Chisti & Moo-Young 1986; Feliu et al. 1998; Taylor et al. 2001; Fykse et al. 2003; Borthwick et al. 2005; Hackley & Wiesner 2010). It would be desirable to replace the traditional sonication with a cell disruption method of greater reproducibility and reduced thermic stress. It

has been shown repeatedly, that sonication can lead to local temperature changes of as much as 90 K close to the sonicator probe and that sufficient and even cooling is challenging (Chisti & Moo-Young 1986; Feliu et al. 1998). The ageing of sonicator probes as experienced during this work underlines the dynamic nature of sonication. While the sonication is somewhat troublesome, alternatives are not readily available. Extensive studies of cell lysis methods for extract preparation have shown that neither bead beating nor high-pressure homogenisation give rise to extracts of higher activity, although results can be somewhat more reproducible depending on the precise equipment and conditions used (Kigawa et al. 2004; Shrestha et al. 2012; Sun et al. 2013). Alternative approaches such as freeze-thaw cycling and lysozyme incubation have similar lysis efficiencies but do not give rise to active cell extracts (Shrestha et al. 2012). It remains to be elucidated if chemical lysis reagents, such as BugBuster, present a viable option in the preparation of cell free expression extracts. It appears likely that chemical lysis would at least require a subsequent dialysis step.

5.4 Conclusion

It has been shown that the preparation of cell free transcription translation extracts from *E. coli* following the methods published by Kwon & Jewett in 2015 gives rise to functional extracts with an activity around 25% of that of commercial S30 extracts from Promega, as estimated using pBest-*luc* luciferase assay. Using T7 based extracts, the aforementioned authors were able to match the activity of commercial T7 extracts. No such data is available for S30 extract systems. Data presented in this work suggests that a run-off reaction, potentially removing inhibitory components from the extracts, may be of greater importance for the S30 extracts than anticipated from published data available for T7 extracts. Despite being widely employed for cell disruption, sonication may not present the ideal lysis method for expression extract production, due to limited reproducibility and difficult to control temperature spikes. A truly practical, streamlined, high-throughput and reproducible method for the preparation of cell free expression extracts is yet to be established.

6 Spinach2 in cell free systems

6.1 Introduction

Cell free expression systems are a powerful tool in applications where cell based expression systems are too variable, slow, difficult to store or prohibitive due to legislation on the release of genetically modified organisms.

However, cell free expression systems are not without their limitations. Their very nature is to be a defined, precisely controlled system. The lack of a self-replicating functionality means that they have to function on a limited supply of intracellular supplies such as ATP sources, amino acids and nucleotides. In consequence the *in vitro* synthesis of proteins is limited in two dimensions: the overall maximum protein amount that can be synthesized with the supplied components, and the time during which the cell free expression system stays active before background processes and chemical deterioration have used up supplies or inactivated the system (Carlson et al. 2012; Bernhard & Tozawa 2013; Kwon & Jewett 2015; Promega Corporation 2015a).

Consequently, the use of cell free expression systems in biosensors that rely on protein expression, while having many advantages, also presents a limitation. Sensors that can function without the need for protein expression reduce the requirements on such a system drastically.

In recent years fluorescent RNA aptamers have been created that mimic fluorescent proteins without the need for translation (Paige et al. 2011). One of these aptamers, an enhanced version of the first such aptamer to be published, is Spinach2 (Strack et al. 2013). Fluorescent RNA aptamers are based on the same principles that allow protein fluorescence. In proteins such as GFP, intramolecular cyclization of AAs leads to the formation of a fluorophore, 4-hydroxybenzlidene imidazolinone (HBI) (Stepanenko et al. 2011).

Contact points between the protein and the fluorophore prevent intramolecular motions of the latter, resulting in fluorescence being the only major way for the fluorophore to dissipate the energy from excitation (Meech 2009). A slight variation of HBI can be used in Spinach2, 3,5-dimethoxy-4-hydroxybenzylidene

imidazolinone (DMHBI). The secondary structure of the aptamer presents a loop in which DMHBI can bind, conferring the same intramolecular immobilisation that confers fluorescence in fluorescent proteins (figure 6.1-1, Paige et al. 2011; Strack et al. 2013). Changing the chemical to (Z)-4-(3,5-difluoro-4-hydroxybenzylidene)-2-methyl-1-(2,2,2-trifluoroethyl)-1H-imidazol-5(4H)-one, or slightly less tongue breaking DFHBI-1T, increases fluorescence and moves the excitation/emission peaks to 482/505 nm, close enough to GFP fluorescence to become measurable with GFP filter sets (Song et al. 2014).

The 95 base core sequence of Spinach2 is usually flanked by a tRNA scaffold, increasing the stability of the Spinach2 secondary structures, leading to a total sequence length of just 168 bases (Ponchon & Dardel 2007; Strack et al. 2013).

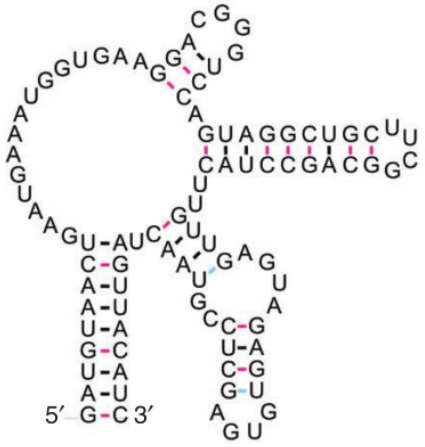

Figure 6.1-1: Secondary structure of Spinach2. Modified from Strack et al. 2013.

In theory, expression of these aptamers requires only a functioning RNA polymerase and hence a drastically reduced supply of salts, energy sources, the fluorescence chemical and other compounds when compared to complete cell free expression systems. Bearing in mind the metal sensing plasmids introduced in chapter 4, a translation free approach might be an alternative reducing the cost and complexity of the sensor. In supplying the regulatory proteins as ingredients in the cell free transcription reaction, only the transcription itself is required from the cell free system.

6.2 Aims of work presented in this chapter

The aims of this chapter were to establish the potential of the fluorescent RNA aptamers, using the example of Spinach2, as next generation reporters in cell free biosensors. The aims in particular were:

- To incorporate Spinach2 into the existing biosensor concept.
- To test *in vivo* transcription of Spinach2 to gain experience with the aptamer.
- To test *in vitro* metal sensing with the Spinach2 metal sensors.
- To identify the minimal requirements for Spinach2 transcription *in vitro* as a first step towards a minimal, high speed cell free sensor approach.

6.3 Results

The Spinach2 sequence including the tRNA scaffold was obtained as custom synthesized (figure 6.3-1), sequence verified dsDNA from Integrated DNA Technologies (Leuven, Belgium) and used in PaperClip assemblies (2.2.1.5) to construct Spinach2 versions of the biosensor constructs based on P_{copA}-WT and P_{zntA}-sRBS as well as the P_{lac} positive control previously introduced, replacing the *sce* gene in these constructs (4.2.2).

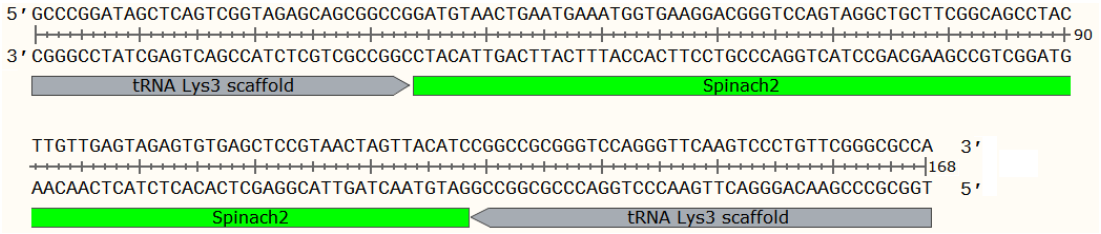

Figure 6.3-1: DNA sequence of Spinach2 and the tRNA Lys2 scaffold.
Total sequence length is given at the end of each line and location of the scaffold and the Spinach2 sequence are indicated by boxes below the sequence. The Spianch2 sequence corresponds to the sequence published by Strack et al. in 2013.

Initial data obtained from the pBest-P_{lac}-Spinach2 based positive control indicated very low fluorescence levels (data not shown). A T7-promoter based approach was to be developed but due to time constraints a pET31b vector containing Spinach2 with its scaffold under the control of the T7 promoter (Wang et al. 2016) was obtained from the Addgene plasmid repository (Cambridge, MA, USA, figure 6.3-2).

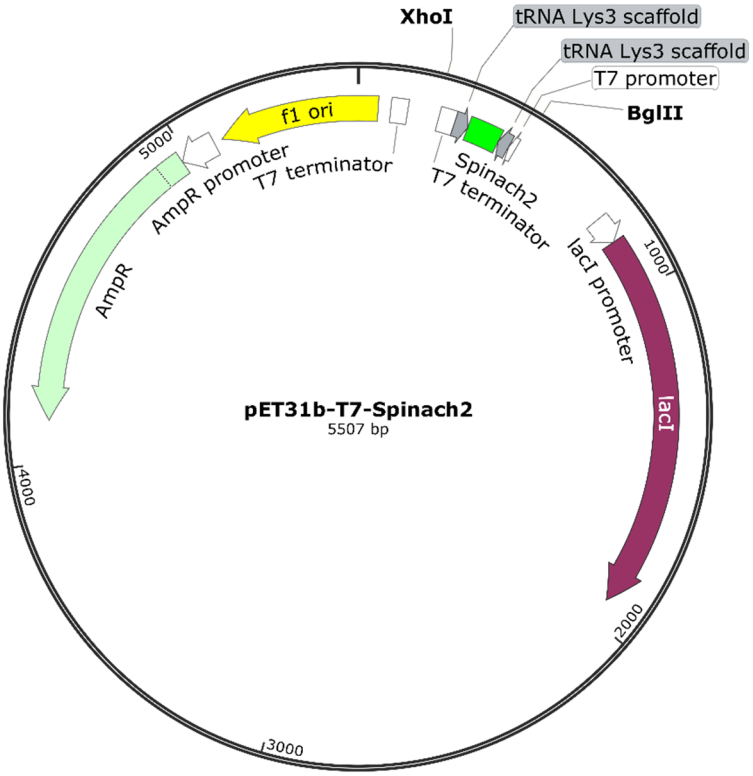

Figure 6.3-2: Plasmid map of pET31b-T7-Spinach2.
Relevant features given in the figure are the ampicillin resistance conferring *AmpR* gene and its promoter, the *lacI* gene and its promoter, the f1 origin of replication and the T7 promoter controlling Spinach2 with its tRNA scaffold, followed by two T7 terminators. Restriction sites are shown for *Xho*I and *Bgl*II (Wang et al. 2016).

6.3.1 *In vivo* screening of Spinach2

All aforementioned constructs were initially tested *in vivo*. According to Pothoulakis et al. *in vivo* fluorescence of Spinach (or in this case Spinach2) can be achieved by incubating an *E. coli* culture expressing Spinach2 in LB medium with DFHBI-1T (Pothoulakis et al. 2013, see 2.2.8 for details). The results were somewhat discouraging as fluorescence of samples containing DFHBI-1T was only marginally higher than fluorescence of samples without the chemical (figure 6.3-3).
To exclude autofluorescence of DFHBI-1T in absence of the Spinach2 aptamer as the cause of fluorescence differences, samples with and without the chemical were independently blanked against LB medium with and without DFHBI-1T. The P_{T7}

positive control did not perform as expected and fluorescence was not found to be stronger than fluorescence of P_{lac} controls (figure 6.3-3).

The DFHBI-1T free control of P_{T7} showed a much greater variation between samples than P_{lac}. Results for P_{copA}-WT samples indicated no enhanced fluorescence in presence of DFHBI1-T over the control. It remained unclear whether this was due to a problem in the controls for this construct, causing higher fluorescence than anticipated. The P_{zntA}-sRBS-Spinach2 plasmid led to the only encouraging result, showing a fluorescence increase in presence of DFHBI-1 greater than the errors of the data. It should be noted that fluorescence was very weak in all samples and gain settings in the instrument had to be set to high values (2835 on a 0 to 4095 scale) to obtain reasonable readings. The consequence of this high gain setting may have been increased background noise.

The experiment was repeated under the same conditions, using a 480 nm excitation filter to reduce crosstalk and with an extended 20 min incubation time with DFHBI-1T. Results remained unchanged in the tendencies described above but overall fluorescence was even weaker (not shown). As DFHBI-1T stocks were depleting and due to time constraints, focus was shifted towards *in vitro* experiments.

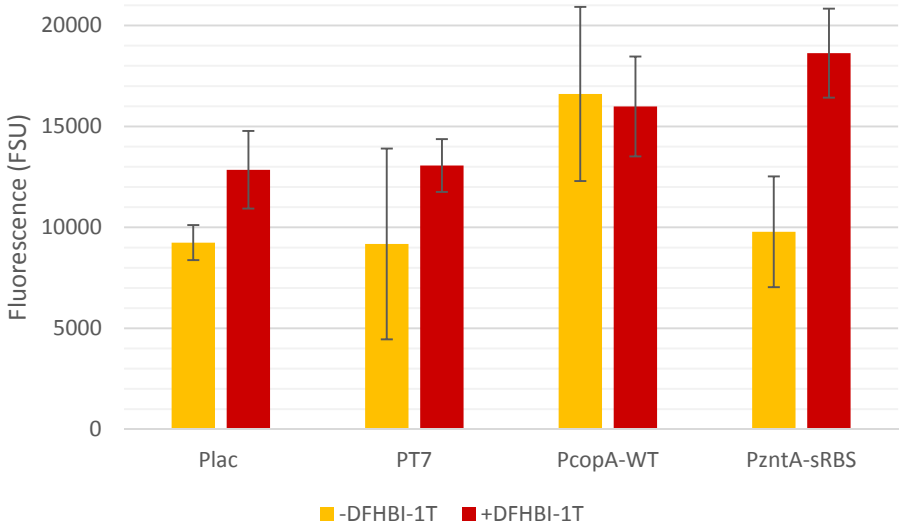

Figure 6.3-3: Spinach2 fluorescence at 485/505 nm *in vivo*.
Data shown represents blank corrected triplicates with one SD above and below average indicated by error bars.

6.3.2 *In vitro* screening of Spinach2 fluorescence in the S30 expression system

Based on the unexpectedly low fluorescence observed for both of the positive controls as well as the high variation for the P_{copA} construct, it was decided to perform *in vitro* experiments only for the P_{zntA}-sRBS plasmid and the P_{lac} control as the S30 expression system (Promega) did not contain T7 RNA polymerase. In relation to the *sce* expression based cell free experiments described previously (4.2.7), influence of Zn, Hg, Cd and Pb on the cell free system expressing Spinach2 and induction of P_{zntA}-sRBS were tested. Concentration ranges were chosen based on the experiences gained with the *sce* system and 10 µL reactions were carried out at 37°C for 4 h in a 384 well low volume plate. DFHBI-1T was added to reactions to a final concentration of 200 µM and plates were incubated on ice for 10 min after addition of the chemical. Fluorescence measurements were performed at RT but great care was taken to minimise time between removal of the plate from the ice and the measurement.

High gain settings had to be used, possibly resulting in some background noise. Water, and S30 reactions without template DNA but with DFHBI-1T as well as

without DFHBI-1T but with template DNA were used as negative controls. Low fluorescence intensities and large errors did not allow for confident interpretation of the results. However, Zn may have had a positive influence on fluorescence intensity and Pb may have had an adverse effect on fluorescence (figure 6.3-4 A). The latter effect became important when analysing the results of the induction testing of P_{zntA}-sRBS (figure 6.3-4 B). Pb at 10 µM concentration led to the strongest increase in any of the P_{zntA}-sRBS samples while the opposite could be expected based on the results from the P_{lac} testing. As results were unsatisfying, *in vitro* testing of the T7 promoter based plasmid was considered as a means of gaining some deeper insight in the remaining time.

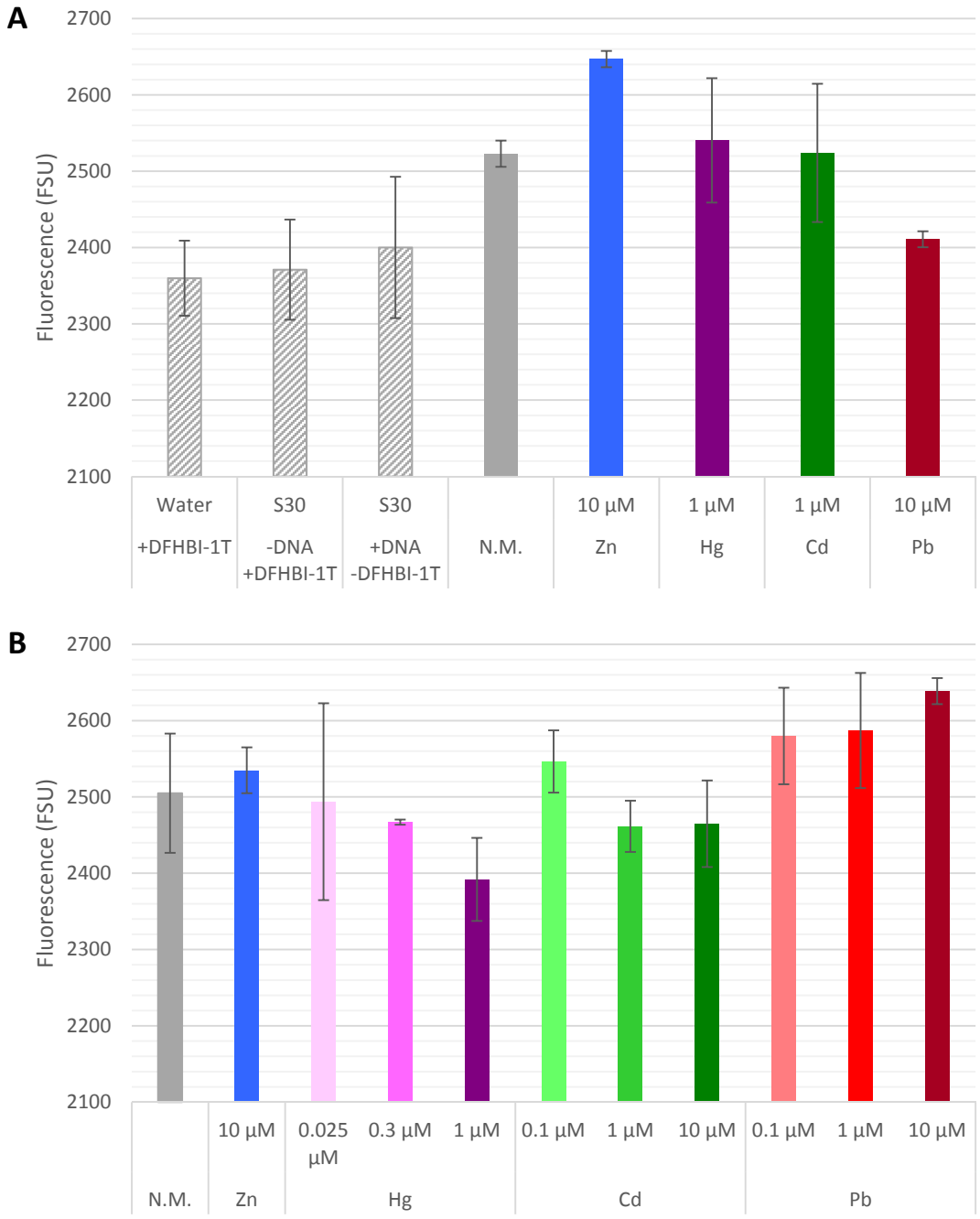

Figure 6.3-4: Spinach2 fluorescence in the S30 expression system.

Data shown represent the average of triplicates with one SD above and below shown as error bars. Note that the Y-axis does not cut at 0. A) Influence of heavy metals on the P_{lac} positive control. Dashed bars show controls with water, S30 reactions without template DNA and reactions with template DNA but no DFHBI-1T. Solid grey bar (N.M) indicates P_{lac} positive control without metal. Controls with added metals are colour coded to part B of the figure. B) P_{zntA}-sRBS controlled fluorescence of Spinach2 in presence of heavy metals. N.M indicates the uninduced (no metal) control.

6.3.3 T7 RNA polymerase based i*n vitro* expression of Spinach2

T7 RNA polymerase (NEB) was used to transcribe the aforementioned pET31b-T7-Spinach2 plasmid (figure 6.3-2). Reactions were carried out according to the manufacturer's recommendations in the supplied buffer (2.2.10). Incubation times of 4 h and 16 h before addition of DFHBI-1T were tested but no difference in Spinach2 fluorescence was observed. Similarly, the use of circular plasmid or linearised DNA as a template in the T7 RNA polymerase reaction did not give rise to significantly changed results. As the pET31b-T7-Spinach2 plasmid is relatively large at 5507 bp, it was digested with *Xho*I and *Bgl*II, resulting in a 255 bp linear fragment incorporating the T7 promoter, Spinach2 with its scaffold and the T7 terminator sequence. The digest was loaded onto an agarose gel and bands with the correct size were cut from the gel and purified (2.2.2.1 and 2.2.2.2). After the T7 RNA polymerase reaction, DFHBI-1T was added to the reactions and the mixture was incubated on ice for 10 min prior to measurement in a 384-well low volume plate at 485/505 nm. As shown in figure 6.3-5 below, fluorescence was relatively uniform between all tested samples, indicating no fluorescence of detectable intensity was emitted by the Spinach2 aptamer.

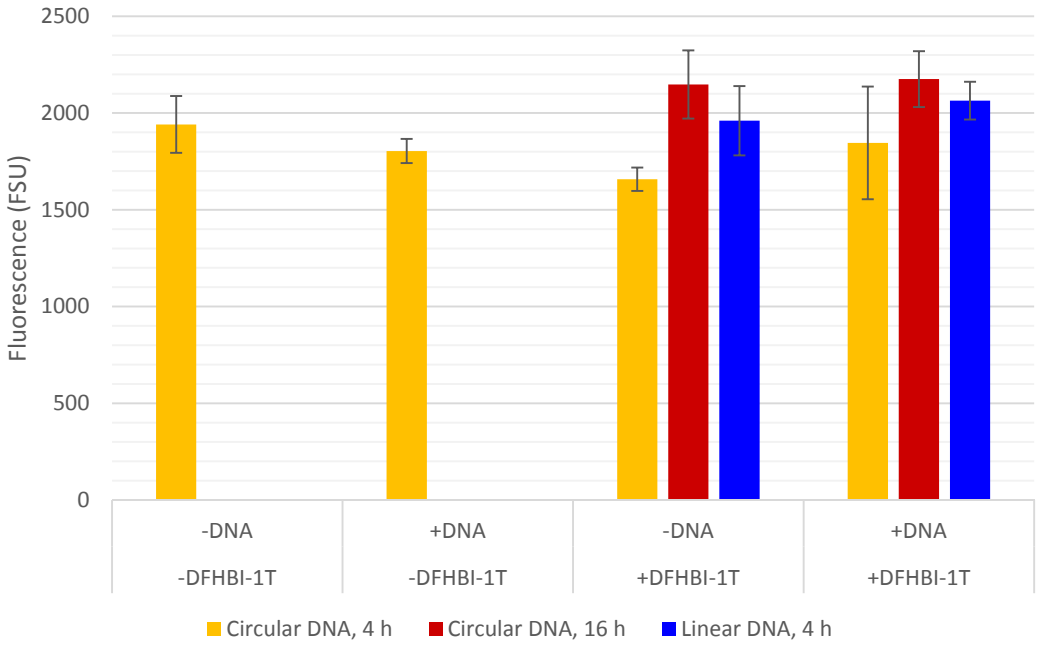

figure 6.3-5: Fluorescence of T7 RNA polymerase reactions transcribing Spinach2. Experiments were carried out with circular template DNA for 4 h and 16 h and with linear template DNA for 4 h. Error bars show one SD above and below the average of triplicates.

To exclude lack of transcription as a possible cause for the absence of Spinach2 fluorescence, reactions were run on 1.5% agarose gels after fluorescence measurements (2.2.2.1). The expected transcript length was 168 bases. All reactions containing template DNA, circular or linear, showed clear bands located between 100 and 200 bp on the included DNA ladder (figure 6.3-6 A and B). Those reactions that had been performed with the full pET plasmid as a template showed an additional band at slightly above 300 bp (figure 6.3-6 A). This size concurs with the expected transcription length if the T7 terminator directly downstream of the Spinach2 sequence in the pET31b-T7-Spinach2 plasmid did not exhibit 100% termination efficiency and some transcription carried on to the second T7 terminator. The latter scenario would give rise to a transcript of approximately 306 bases. Due to the clear bands visible in all samples that had been incubated with template DNA, regardless of DFHBI-1T presence, it was concluded that transcription was taking place and that the lack of Spinach2 fluorescence was caused by other factors. Due to time constraints no further investigations could be made.

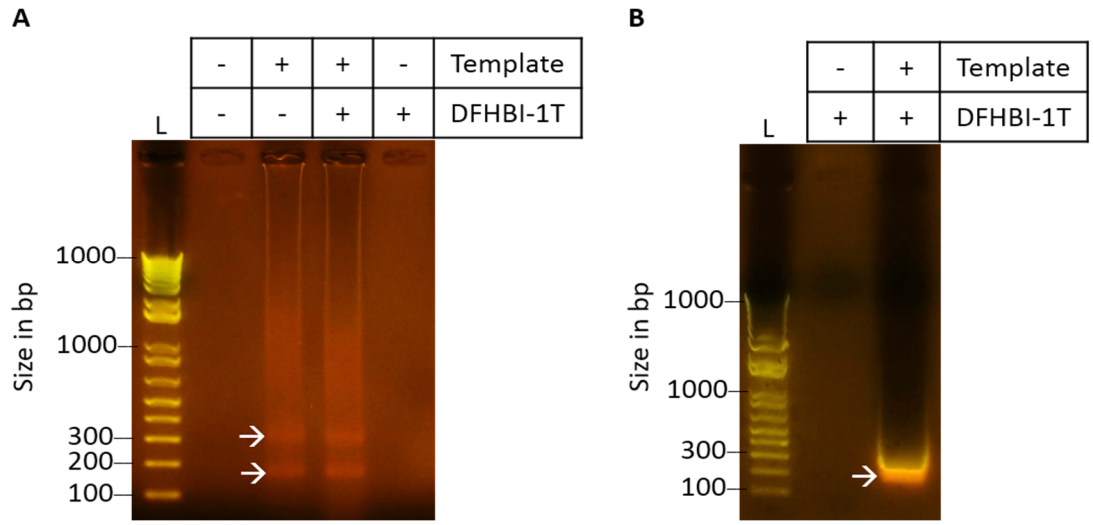

Figure 6.3-6: Agarose gel images of T7 RNA polymerase reactions.

1.5% Agarose gels were stained with SafeView nucleic acid stain (NBS biologicals, Cambridgeshire, UK) and visualised under blue light. Relevant sizes of the 1kb plus DNA ladder (Invitrogen) are indicated. Sample bands likely representing the RNA synthesized by the T7 RNA polymerase in the supplied NTP buffer are indicated by arrows (→). Presence (+) or absence (-) of template DNA and DFHBI-1T in the T7 RNA polymerase reactions are indicated above the gel images. A) Circular DNA template pET31b-T7-Spinach2. B) Linear 168 bp DNA template consisting of Spinach2 with its tRNA scaffold.

6.4 Discussion

Spinach and Spinach2 have been shown to be practical replacements for fluorescent proteins where the latter are prohibitive due to size, the need for translation or the dependency on oxygen (Shaner et al. 2005; Paige et al. 2011; Strack et al. 2013). Recently their use in biosensors for the anaerobic detection of cyclic di-GMP has been reported, proving further their versatility and functionality (Wang et al. 2016).

As elaborated previously (6.1) the inclusion of Spinach2 or a similar fluorescent RNA aptamer in the cell free metal biosensors would have been crucial for the further simplification of the presented biosensors. *In vivo* results may indicate a weak fluorescence in the P_{zntA}-sRBS samples but later *in vitro* experiments make this highly speculative. A range of causes for the lack of fluorescence is possible: lack of transcription, expression levels too low to detect sufficiently strong fluorescence, incorrect folding of the aptamer, rapid degradation of Spinach2, inadequate DFHBI-1T supply or deterioration of the same, insufficient penetration of cells by DFHBI-1T and lastly technical issues in the plate reader employed as well as experimental errors.

Transcription has been shown *in vitro* and since P_{lac}, P_{copA} and P_{zntA} constructs are identical in their promoter configuration to the *sce* based sensors, it appears unlikely that a total lack of transcription is the cause of the unsatisfying results. Similarly, expression would be expected to be strong enough to detect some fluorescence, if not *in vivo* or in the S30 expression system, then in the T7 RNA polymerase experiments. Only assessment of the actual RNA levels for the Spinach2 aptamer in the cell or each *in vitro* reaction would deliver conclusive results in this regard.

Correct folding of Spinach2 has not been reported to be of an unusually challenging nature (Strack et al. 2013; Wang et al. 2016). The agarose gels presented indicate presence of nucleic acid chains, making degradation another unlikely cause for the lack of fluorescence. If experiments were to be repeated employing an RNA ladder and under conditions ensuring the suppression of secondary structures, the correct size of the transcripts could be confirmed confidently. As a result of the exclusion of the aforementioned causes, inadequacies of the DFHBI-1T supplied appear to be a more possible cause. DFHBI-1T was commercially available from two suppliers at

the time of writing and had been supplied by Tocris Bioscience. Solutions used throughout this work were prepared freshly at regular intervals, stored in the dark at -20°C and not used if older than 7 days. Ideally the quality of the DFHBI-1T stocks would be assessed by chemical analysis, a procedure beyond the scope of this work. Repetition of the experiments with a new batch of DFHBI-1T would appear to be the most sensible next step in these experiments, reducing the likelihood of substrate inadequacies or experimental errors causing the unsatisfactory results.

As the 485 nm excitation and 505 nm emission filters employed for measuring Spinach2 fluorescence had been used for the measurement of GFP fluorescence in another project within the institute, it appears unlikely that technical issues had any influence on the results. It is however possible, that analysis of experiments with more sensitive devices might have led to different results.

Interestingly, analysis of the T7 RNA polymerase reactions on agarose gels has indicated similar amounts of Spinach2 aptamer are transcribed in 4 and 16 h reactions and with circular or linear DNA templates. This may be indicative of shorter incubation times, such as 1 h or below, being sufficient for abundant transcription of Spinach2. If satisfactory fluorescence levels could be achieved in future experiments, cell free transcription without the need for a complete cell extract based expression system would present a drastic simplification of the cell free metal sensors with potential benefits in regard to cost, speed, reliability and ultimately practicality.

6.5 Conclusion

Spinach2, a fluorescent RNA aptamer mimicking the green fluorescent protein GFP might present the next step in simplifying the previously presented cell free metal sensors (4). Sensors not requiring any translation could exhibit superior detection speeds, increased reliability and a reduced cost per reaction. A range of experiments has been carried out to demonstrate Spinach2 fluorescence *in vivo* and *in vitro* but no clear results could be obtained. Available data indicates some weak fluorescence but further experiments are required. Data presented indicates that simple T7 RNA polymerase reactions may be sufficient for the transcription of the Spinach2 aptamer.

7 Conclusion

The presented work focussed on the development of biosensors for heavy metals. When reviewing the requirements for field-use sensors, it becomes apparent that simplicity of use and design are crucial. With AsGard a solid plastic support biosensor has been presented which holds the potential for true field applicability if visibility and robustness can be improved. In its current state it already detects As(III) at relevant concentrations. Interestingly the necessary fusion of mCherry to the ArsR protein has shown that C-terminal fusions to the latter protein influence its As(III) responsiveness. This property, corresponding to available literature (Merulla et al. 2013), has been shown again in the cell free transcription-translation sensors presented in chapter 4. While many current metal biosensor designs employing regulatory proteins are based on the expression of these proteins within the sensor (1.2), the designs presented throughout this work rely on the supply of these proteins to the sensors in a more or less purified form. While possibly increasing the manufacturing cost of a final sensors, this strategy unlocks a range of potential innovations for cell free biosensor design exemplified in this work. The supply of precise amounts of regulatory proteins allows for tuneable sensor characteristics as exemplified with the As-sensors based on ArsR controlling P_{ars} and sensing As(III) at relevant concentrations. Previously such systems relied on the variation of promoter strength ratios controlling the regulatory proteins, entailing the necessity for an individual set of plasmids for each adjustment desired and making a precise control challenging (Merulla et al. 2013). Additional sensors presented on the basis of P_{copA} and P_{zntA} may be tuneable in a similar fashion to the P_{ars} based sensor in the future and in their current state show detection ranges close to the desired thresholds for the assessment of drinking water quality, as defined by various national and international bodies. Approaches to further reduce the sensor complexity of cell free sensors to fully eliminate the need for translation have been presented (chapter 6). If current problems can be overcome and these approaches can be developed into working transcription-only sensors, possibly utilising fluorescent RNA aptamers, a simple multiplexed sensor with the ability to detect a wide range of heavy metals comes within reach. Such a sensor could have the potential to complement current water analysis techniques.

8 References

Abrevaya, X.C., Sacco, N.J., Bonetto, M.C., Hilding-Ohlsson, A. & Cortón, E., 2015. Analytical applications of microbial fuel cells. Part II: Toxicity, microbial activity and quantification, single analyte detection and other uses. *Biosensors and Bioelectronics*, 63, pp.591–601.

Adriano, D.C., 2001. Cadmium. In *Trace elements in terrestrial environments*.

Agency for Toxic Substances and Disease Registry, 2007. *Toxicological Profile for Arsenic*, Atlanta.

Aleksic, J., Bizzari, F., Cai, Y., Davidson, B., Mora, K. De, Ivakhno, S., Seshasayee, S.L., Nicholson, J., Wilson, J., Elfick, A., French, C., Ma, H. & Millar, A., 2007. Development of a novel biosensor for the detection of arsenic in drinking water. *IET Synth. Biol.*, 1, pp.87–90.

Alieva, N.O., Konzen, K. a, Field, S.F., Meleshkevitch, E. a, Hunt, M.E., Beltran-Ramirez, V., Miller, D.J., Wiedenmann, J., Salih, A. & Matz, M. V, 2008. Diversity and evolution of coral fluorescent proteins. *PloS one*, 3(7), p.e2680.

Anon, 2015. Bielefeld-CeBiTec. Available at: http://2015.igem.org/Team:Bielefeld-CeBiTec.

Baker, D. & Gough, D., 1996. Dynamic delay and maximal dynamic error in continuous biosensors. *Analytical chemistry*, 68(8), pp.1292–1297.

Bakir, F.E.A., 1973. Methylmercury Poisoning in Iraq. *Science*, 181, pp.230–241.

Bell, E.A., Boehnke, P., Harrison, T.M. & Mao, W.L., 2015. Potentially biogenic carbon preserved in a 4.1 billion-year-old zircon. *Proceedings of the National Academy of Sciences of the United States of America*, 112(47), pp.14518–21.

Bereza-Malcolm, L.T., Mann, G. & Franks, A.E., 2015. Environmental Sensing of Heavy Metals Through Whole Cell Microbial Biosensors: A Synthetic Biology Approach. *ACS Synthetic Biology*, 4(5), pp.535–546.

Berg, M., Tran, H., Nguyen, T., Pham, H., Schertenleib, R. & Giger, W., 2001. Arsenic contamination of groundwater and drinking water in Vietnam: a human health threat. *Environmental Science and Technology*, 35(13), pp.2621–2626.

Berman, J., Eisenberg, S. & Tye, B., 1987. An Agarose Gel Electrophoresis Assay for the Detection of DNA-Binding Activities in Yeast Cell Extracts. *Methods in Enzymology*, 155(1981), pp.528–537.

Bernhard, F. & Tozawa, Y., 2013. Cell-free expression-making a mark. *Current Opinion in Structural Biology*, 23(3), pp.374–380.

Bertani, G., 1951. Studies on lysogenesis. I. The mode of phage liberation by lysogenic Escherichia coli. *Journal of bacteriology*, 62(3), pp.293–300.

Binet, M.R.B. & Poole, R.K., 2000. Cd (II), Pb (II) and Zn (II) ions regulate expression of the metal-transporting P-type ATPase ZntA in Escherichia coli. *FEBS Letters*, 473(1), pp.67–70.

Bjerketorp, J., Håkansson, S., Belkin, S. & Jansson, J.K., 2006. Advances in

preservation methods: Keeping biosensor microorganisms alive and active. *Current Opinion in Biotechnology*, 17(1), pp.43–49.

Bobade, S., Kalorey, D.R. & Warke, S., 2016. Biosensor Devices: A review on their biological applications. *Bioscience Biotechnology Research Communications*, 9(1), pp.132–137.

Bonnet, J., Yin, P., Ortiz, M.E., Subsoontorn, P. & Endy, D., 2013. Amplifying Genetic Logic Gates. *Science*, 340(May), pp.599–603.

Borthwick, K.A.J., Coakley, W.T., McDonnell, M.B., Nowotny, H., Benes, E. & Gröschl, M., 2005. Development of a novel compact sonicator for cell disruption. *Journal of Microbiological Methods*, 60(2), pp.207–216.

Bousse, L., 1996. Whole cell biosensors. *Sensors and Actuators B: Chemical*, B34, pp.270–275.

Bradford, M.M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*, 72(1–2), pp.248–254.

Bradl, H., Kim, C., Kramar, U. & StÜben, D., 2005. Heavy Metals in the Environment: Origin, Interaction and Remediation H. B. Bradl, ed. *Interface Science and Technology*, 6, pp.28–164.

Bradl, H. & Xenidis, A., 2005. Chapter 3 Remediation techniques. *Interface Science and Technology*, 6(C), pp.165–261.

Bradl, H.B. ed., 2005. *Heavy metals in the environment* First Edit., London: Elsevier Ltd.

Brännvall, M., Bindler, R. & Renberg, I., 1999. The Medieval metal industry was the cradle of modern large-scale atmospheric lead pollution in northern Europe. *science & technology*.

Brocklehurst, K.R., Hobman, J.L., Lawley, B., Blank, L., Marshall, S.J., Brown, N.L. & Morby, A.P., 1999. ZntR is a Zn(II)-responsive MerR-like transcriptional regulator of zntA in Escherichia coli. *Molecular Microbiology*, 31(3), pp.893–902.

Brown, N.L., Stoyanov, J. V, Kidd, S.P. & Hobman, J.L., 2003. The MerR family of transcriptional regulators. *FEMS Microbiology Reviews*, 27(2–3), pp.145–163.

Budak, G.G., Ozkan, C.S. & Ozkan, M., 2013. Nanomedicine and the Nose. In *Nasal Physiology and Pathophysiology of Nasal Disorders*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 589–597.

Busenlehner, L.S., Pennella, M.A. & Giedroc, D.P., 2003. The SmtB/ArsR family of metalloregulatory transcriptional repressors: Structural insights into prokaryotic metal resistance. *FEMS Microbiology Reviews*, 27(2–3), pp.131–143.

Callender, E., 2003. Heavy metals in the environment-historical trends. *Treatise on geochemistry*, 9, pp.67–105.

Campbell, D.R., Chapman, K.E., Waldron, K.J., Tottey, S., Kendall, S., Cavallaro, G., Andreini, C., Hinds, J., Stoker, N.G., Robinson, N.J. & Cavet, J.S., 2007. Mycobacterial cells have dual nickel-cobalt sensors: sequence relationships and

metal sites of metal-responsive repressors are not congruent. *The Journal of biological chemistry*, 282(44), pp.32298–310.

Carlin, A., Shi, W., Dey, S. & Rosen, B.P., 1995. The ars operon of Escherichia coli confers arsenical and antimonial resistance. *Journal of Bacteriology*, 177(4), pp.981–986.

Carlson, E.D., Gan, R., Hodgman, C.E. & Jewett, M.C., 2012. Cell-free protein synthesis: Applications come of age. *Biotechnology Advances*, 30(5), pp.1185–1194.

Castillo, J., Gáspár, S., Leth, S., Niculescu, M., Mortari, A., Bontidean, I., Soukharev, V., Dorneanu, S.A., Ryabov, A.D. & Csöregi, E., 2004. Biosensors for life quality - Design, development and applications. *Sensors and Actuators, B: Chemical*, 102(2), pp.179–194.

Cebrián, M.E., Albores, A., Aguilar, M. & Blakely, E., 1983. Chronic arsenic poisoning in the north of Mexico. *Human toxicology*, 2(1), pp.121–33.

Changela, A., Chen, K., Xue, Y., Holschen, J., Outten, C.E., O'Halloran, T. V & Mondragón, A., 2003. Molecular Basis of Metal-Ion Selectivity and Zeptomolar Sensitivity by CueR. *Science*, 301(September), pp.1383–1388.

Chen, J., Chen, Y., Liu, W., Bai, C., Liu, X., Liu, K., Li, R., Zhu, J.H. & Huang, C., 2012. Developmental lead acetate exposure induces embryonic toxicity and memory deficit in adult zebrafish. *Neurotoxicology and Teratology*, 34(6), pp.581–586.

Chen, J. & Rosen, B., 2014. Biosensors for Inorganic and Organic Arsenicals. *Biosensors*, 4(4), pp.494–512.

Chen, P. & He, C., 2003. A general strategy to convert the MerR family proteins into highly sensitive and selective flourescent biosenors for metal ions. *Journal of American Chemical Societies*, 126(3), pp.728–729.

Chisti, Y. & Moo-Young, M., 1986. Disruption of microbial cells for intracellular products. *Enzyme and Microbial Technology*, 8(4), pp.194–204.

Christie, G.E., White, T.J. & Goodwin, T.S., 1994. Amer Rhomologue at 74 minutes on the Escherichia coli genome. *Gene*, 146(1), pp.131–132.

Chung, C.T., Niemela, S.L. & Miller, R.H., 1989. One-step preparation of competent Escherichia coli: Transformation and storage of bacterial cells in the same solution (recombinant DNA). *Pnas*, 86(April), pp.2172–2175.

Clark, K.L., Halay, E.D., Lai, E. & Burley, S.K., 1993. Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. *Nature*, 364(6436), pp.412–420.

Clark, L.C. & Lyons, C., 1962. Electrode systems for continuous monitoring in cardiovascular surgery. *Annals Of The New York Academy Of Sciences*, 102(1), pp.29–45.

Cook, W.J., Kar, S.R., Taylor, K.B. & Hall, L.M., 1998. Crystal structure of the cyanobacterial metallothionein repressor SmtB: a model for metalloregulatory proteins. *Journal of Molecular Biology*, 275(2), pp.337–346.

Costa-Silva, T.A., Nogueira, M.A., Fernandes Souza, C.R., Oliveira, W.P., Said, S., Oliveira, W.P. & Said, S., 2011. Lipase Production by Endophytic Fungus Cercospora kikuchii: Stability of Enzymatic Activity after Spray Drying in the Presence of Carbohydrates. *Drying Technology*, 29(9), pp.1112–1119.

Couñago, R.M., Chen, N.H., Chang, C.-W., Djoko, K.Y., McEwan, A.G. & Kobe, B., 2016. Structural basis of thiol-based regulation of formaldehyde detoxification in H. influenzae by a MerR regulator with no sensor region. *Nucleic Acids Research*, 44(14), pp.6981–93.

Crawford, G.M. & Tavares, O., 1974. Simple hydrogen sulfide trap for the Gutzeit arsenic determination. *Analytical Chemistry*, 46(July), p.1974.

Date, A., Pasini, P. & Daunert, S., 2007. Construction of spores for portable bacterial whole-cell biosensing systems. *Analytical Chemistry*, 79(24), pp.9391–9397.

Date, A., Pasini, P., Sangal, A. & Daunert, S., 2010. Packaging sensing cells in spores for long-term preservation of sensors: A tool for biomedical and environmental analysis. *Analytical Chemistry*, 82(14), pp.6098–6103.

Daunert, S., Barrett, G., Feliciano, J.S., Shetty, R.S., Shrestha, S. & Smith-Spencer, W., 2000. Genetically engineered whole-cell sensing systems: coupling biological recognition with reporter genes. *Chemical Reviews*, 100(7), pp.2705–2738.

Day, R.N. & Davidson, M.W., 2009. The fluorescent protein palette: tools for cellular imaging. *Chemical Society reviews*, 38(10), pp.2887–921.

Dey, M.K., Satpati, A.K. & Reddy, A.V.R., 2014. Electrodeposited antimony and antimony–gold nanocomposite modified carbon paste electrodes for the determination of heavy metal ions. *Analytical Methods*, 6(14), p.5207.

Duffus, J.H., 2002. *"Heavy metals" - A meaningless term? (IUPAC Technical Report)*,

Dunlap, W.H., Watson, J.G. & Hume, D.M., 1974. Anemia and Neutropenia Caused by Copper Deficiency. *Annals of Internal Medicine*, 80(4), pp.470–6.

Durrieu, C. & Tran-Minh, C., 2002. Optical Algal Biosensor using Alkaline Phosphatase for Determination of Heavy Metals. *Ecotoxicology and Environmental Safety*, 51(3), pp.206–209.

Duruibe, J.O., Ogwuegbu, M.O.C. & Egwurugwu, J.N., 2007. Heavy metal pollution and human biotoxic effects. *International Journal of Physical Sciences*, 2(5), pp.112–118.

Eisler, R., 1988. Arsenic Hazards to Fish, Wildlife, and Invertebrates: A Synoptic Review. *Contaminant Hazard Reviews*, 85(12), pp.1–38.

El-Nekeety, A., El-Kady, A., Soliman, M., Hassan, N. & Abdel-Wahhab, M., 2009. Protective effect of Aquilegia vulgaris (L.) against lead acetate-induced oxidative stress in rats. *Food and Chemical Toxicology*, 47(9), pp.2209–2215.

Endo, S.B.I., Hu, S.E.W., Nielsen, B.M.J., Tsao, G.S.G.T., Zeng, R.U.A. & Zhou, J.Z.W., 2010. *Advances in Biochemical Engineering* 1st ed. T. Scheper, ed., Berlin Heidelberg: Springer.

Environmental Protection Agency, 2016. Table of Regulated Drinking Water Contaminants. , pp.1–9. Available at: https://www.epa.gov/ground-water-and-drinking-water/table-regulated-drinking-water-contaminants [Accessed July 31, 2016].

Ercal, N., Gurer-Orhan, H. & Aykin-Burns, N., 2001. Toxic metals and oxidative stress part I: Mechanisms involved in metal induced oxidative damage. *Current Topics in Medicinal Chemistry*, 1, pp.529–539.

Espah Borujeni, A., Channarasappa, A.S. & Salis, H.M., 2014. Translation rate is controlled by coupled trade-offs between site accessibility, selective RNA unfolding and sliding at upstream standby sites. *Nucleic Acids Research*, 42(4), pp.2646–2659.

European Commission, 2001. *Council Directive 2001/18/EC*, European Union.

Feliu, J.X., Cubarsi, R. & Villaverde, A., 1998. Optimized release of recombinant proteins by ultrasonication of E. coli cells. *Biotechnology and Bioengineering*, 58(5), pp.536–540.

Flegal, A.R. & Smith, D.R., 1992. Current needs for increased accuracy and precision in measurements of low levels of lead in blood. *Environmental Research*, 58(2), pp.125–133.

Francisco, M.S. & Hope, C., 1990. Identification of the metalloregulatory element of the plasmid-encoded arsenical resistance operon. *Nucleic acids ...*, 18(3), pp.619–624.

Frantz, B. & O'Halloran, T. V., 1990. DNA distortion accompanies transcriptional activation by the metal-responsive gene-regulatory protein MerR. *Biochemistry*, 29(20), pp.4747–4751.

French, C.E., Horsfall, L., Barnard, D.K., Duedu, K., Fletcher, E., Joshi, N., Kane, S.D., Lakhundi, S.S., Liu, C.-K., Oltmanns, J., Radford, D., Salinas, A., White, J. & Elfick, A., 2015. Beyond Genetic Engineering: Technical Capabilities in the Application Fields of Biocatalysis and Biosensors. In B. Giese, ed. *Synthetic Biology*. Springer International Publishing, pp. 113–137.

Frey, J., 2007. Biological safety concepts of genetically modified live bacterial vaccines. *Vaccine*, 25(30), pp.5598–5605.

Fridberg, L., Elinder, C.G. & Kjellstroem, T., 1992. *Cadmium. Environmental Health Criteriia 134*, Geneva: World Health Organization, IPCS.

Fried, M.G., 1989. Measurement of protein-DNA interaction parameters by electrophoresis mobility shift assay. *Electrophoresis*, 10(5–6), pp.366–376.

Fried, M.G. & Bromberg, J.L., 1997. Factors that affect the stability of protein-DNA complexes during gel electrophoresis. *Electrophoresis*, 18(1), pp.6–11.

Fried, M.G. & Liu, G., 1994. Molecular sequestration stabilizes CAP-DNA complexes during polyacrylamide gel electrophoresis. *Nucleic acids research*, 22(23), pp.5054–5059.

Fykse, E.M., Olsen, J.S. & Skogan, G., 2003. Application of sonication to release DNA from Bacillus cereus for quantitative detection by real-time PCR. *Journal of Microbiological Methods*, 55(1), pp.1–10.

Gibson, D.G., Benders, G.A., Andrews-Pfannkoch, C., Denisova, E.A., Baden-Tillson, H., Zaveri, J., Stockwell, T.B., Brownley, A., Thomas, D.W., Algire, M.A., Merryman, C., Young, L., Noskov, V.N., Glass, J.I., Venter, J.C., Hutchison Iii, C.A. & Smith, H.O., 2008. Complete Chemical Synthesis, Assembly, and Cloning of a Mycoplasma genitalium Genome. *Science*, 319(5867), pp.1215–1220.

Gibson, R.S. & Scythes, C.A., 1984. Chromium, selenium, and other trace element intakes of a selected sample of Canadian premenopausal women. *Biological Trace Element Research*, 6(2), pp.105–116.

Giedroc, D.P. & Arunkumar, A.I., 2007. Metal sensor proteins: nature's metalloregulated allosteric switches. *Dalton Transactions*, (29), pp.3107–20.

Gladysheva, T.B., Oden, K.L. & Rosen, B.P., 1994. Properties of the arsenate reductase of plasmid R773. *Biochemistry*, 33(23), pp.7288–93.

Guilbault, G.G. & Lubrano, G.J., 1973. An enzyme electrode for the amperometric determination of glucose. *Analytica Chimica Acta*, 64(3), pp.439–455.

Hach, 2013. Arsenic Low Range Test Kit. Available at: http://www.hach.com/arsenic-low-range-test-kit/product?id=7640217303 [Accessed September 29, 2013].

Hackley, V.A. & Wiesner, M.R., 2010. *Ceint / Nist Protocol for Preparation of Nanoparticle Dispersions From Powdered Material Using Ultrasonic Disruption*,

Han, F.X., Banin, A., Su, Y., Monts, D.L., Plodinec, M.J., Kingery, W.L. & Triplett, G.E., 2002. Industrial age anthropogenic inputs of heavy metals into the pedosphere. *Naturwissenschaften*, 89(11), pp.497–504.

Harada, M., 1995. Minamata disease: methylmercury poisoning in Japan caused by environmental pollution. *Critical reviews in toxicology*, 25(1), pp.1–24.

Harms, H., Wells, M.C. & Van Der Meer, J.R., 2006. Whole-cell living biosensors - Are they ready for environmental application? *Applied Microbiology and Biotechnology*, 70(3), pp.273–280.

Harvie, D.R., Andreini, C., Cavallaro, G., Meng, W., Connolly, B.A., Yoshida, K.I., Fujita, Y., Harwood, C.R., Radford, D.S., Tottey, S., Cavet, J.S. & Robinson, N.J., 2006. Predicting metals sensed by ArsR-SmtB repressors: Allosteric interference by a non-effector metal. *Molecular Microbiology*, 59(4), pp.1341–1356.

Hays, H.C.W., Millner, P.A., Jones, J.K. & Rayner-Brandes, M.H., 2005. A novel and convenient self-drying system for bacterial preservation. *Journal of Microbiological Methods*, 63(1), pp.29–35.

Health Canada, 2012. Guidelines for Canadian Drinking Water Quality Summary Table Prepared by the Federal-Provincial-Territorial Committee on Drinking Water of the Federal-Provincial-Territorial Committee on Health and the Environment March 2006. *Environments*, (October 2014), pp.1–16.

Heldwein, E.E. & Brennan, R.G., 2001. Crystal structure of the transcription activator BmrR bound to DNA and a drug. *Nature*, 409(6818), pp.378–382.

Hellman, L.M. & Fried, M.G., 2007. Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nature protocols*, 2(8), pp.1849–61.

Helmann, J.D., Ballard, B.T. & Walsh, C.T., 1990. The MerR metalloregulatory protein binds mercuric ion as a tricoordinate, metal-bridged dimer. *Science*, 247(4945), pp.946–948.

Hill, W.R. & Pillsbury, D.M., 1939. *Argyria: the pharmacology of silver*, Baltimore: The Williams & Wilkins Company.

Hillson, N.J., The SLIC, Gibson, CPEC and SLiCE assembly methods (and GeneArt® Seamless, In-Fusion® Cloning). *j5: automated DNA assembly software manual and how-to guide*, p.6. Available at: https://j5.jbei.org/j5manual/pages/22.html [Accessed July 4, 2016].

HKSAR Government, 2015. *Lead in Drinking Water Incidents*, London: Drinking Water Inspectorate.

Hobman, J.L., Wilkie, J. & Brown, N.L., 2005. A design for life: Prokaryotic metal-binding MerR family regulators. *BioMetals*, 18(4), pp.429–36.

Holler, F.J., Skoog, D.A. & West, D.M., 1996. *Fundamentals of analytical chemistry*, Philadelphia: Saunders College Pub.

Horikoshi, K. & Grant, W.D. (William D.., 1998. *Extremophiles: microbial life in extreme environments* 1st ed. K. Horikoshi & W. Grant, eds., New York: Wiley-Liss.

Hutchison, C.A., Chuang, R.-Y., Noskov, V.N., Assad-Garcia, N., Deerinck, T.J., Ellisman, M.H., Gill, J., Kannan, K., Karas, B.J., Ma, L., Pelletier, J.F., Qi, Z.-Q., Richter, R.A., Strychalski, E.A., Sun, L., Tsvetanova, B., Smith, H.O., Gibson, D.G. & Venter, J.C., 2016. Design and synthesis of a minimal bacterial genome. *Science*, 351(6280), p.aad6253-1ff.

Jakariya, M., Vahter, M., Rahman, M., Wahed, M.A., Hore, S.K., Bhattacharya, P., Jacks, G. & Persson, L.A., 2007. Screening of arsenic in tubewell water with field test kits: evaluation of the method from public health perspective. *The Science of the total environment*, 379(2–3), pp.167–75.

Jing, D., Agnew, J., Patton, W.F., Hendrickson, J. & Beechem, J.M., 2003. A sensitive two-color electrophoretic mobility shift assay for detecting both nucleic acids and protein in gels. In *Proteomics*. pp. 1172–1180.

Johansson, B., 1972. Agarose gel electrophoresis. *Scandinavian Journal of Clinical & Laboratory Investigation*, 29(s124), pp.7–19.

Joshi, N., Wang, X., Montgomery, L., Elfick, a. & French, C.E., 2009. Novel approaches to biosensors for detection of arsenic in drinking water. *Desalination*, 248(1–3), pp.517–523.

Katzen, F., Peterson, T.C. & Kudlicki, W., 2009. Membrane protein expression: no cells required. *Trends in Biotechnology*, 27(8), pp.455–460.

Kaur, H., Kumar, R., Babu, J.N. & Mittal, S., 2015. Advances in arsenic biosensor development - a comprehensive review. *Biosensors and Bioelectronics*, 63, pp.533–545.

Kaur, P. & Rosen, B., 1992. Plasmid-encoded resistance to arsenic and antimony. *Plasmid*, 27(1), pp.29–40.

Kawakami, Y., Siddiki, M.S.R., Inoue, K., Otabayashi, H., Yoshida, K., Ueda, S., Miyasaka, H. & Maeda, I., 2010. Application of fluorescent protein-tagged trans factors and immobilized cis elements to monitoring of toxic metals based on in vitro protein-DNA interactions. *Biosensors & bioelectronics*, 26(4), pp.1466–73.

Keating, M.H., Beauregard, D., Benjey, W.G., Driver, L., Maxwell, W.H., Peters, W.D. & Pope, a a, 1998. *Mercury Study Report to Congress Volume II: An Inventory of Anthropogenic Mercury Emissions in the United States*,

Keen, C. & Gershwin, M., 1990. Zinc deficiency and immune function. *Annual review of nutrition*, 10, pp.415–31.

Kelley, L. & Sternberg, M., 2009. Protein structure prediction on the Web: a case study using the Phyre server. *Nature protocols*, 4(3), pp.363–371.

Khan, S., Brocklehurst, K.R., Jones, G.W. & Morby, A.P., 2002. The functional analysis of directed amino-acid alterations in ZntR from Escherichia coli. *Biochemical and Biophysical Research Communications*, 299(3), pp.438–445.

Kigawa, T., Yabuki, T., Matsuda, N., Matsuda, T., Nakajima, R., Tanaka, A. & Yokoyama, S., 2004. Preparation of Escherichia coli cell extract for highly productive cell-free protein expression. *Journal of Structural and Functional Genomics*, 5(1–2), pp.63–68.

Ko Ferrigno, P., Köhler, G., Milstein, C., Bradbury, A., Plückthun, A., Ellington, A.D., Szostak, J.W., Tuerk, C., Gold, L., Lofblom, J., Feldwisch, J., Tolmachev, V., Carlsson, J., Stahl, S., Frejd, F.Y., Amstutz, P., Koch, H., Binz, H.K. & Deuber, S.A., 2016. Non-antibody protein-based biosensors. *Essays in biochemistry*, 60(1), pp.19–25.

Kulkarni, R.D. & Summers, A.O., 1999. MerR Cross-Links to the α, β, and σ 70 Subunits of RNA Polymerase in the Preinitiation Complex at the merTPCAD Promoter. *Biochemistry*, 38(11), pp.3362–3368.

Kumaraswami, M., Newberry, K.J. & Brennan, R.G., 2010. Conformational Plasticity of the Coiled-Coil Domain of BmrR Is Required for bmr Operator Binding: The Structure of Unliganded BmrR. *Journal of Molecular Biology*, 398(2), pp.264–275.

Kwon, Y.-C. & Jewett, M.C., 2015. High-throughput preparation methods of crude extract for robust cell-free protein synthesis. *Scientific Reports*, 5, p.8663.

Laemmli, U., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 227(August), pp.680–685.

Lehmann, M., Riedel, K., Adler, K. & Kunze, G., 2000. Amperometric measurement of copper ions with a deputy substrate using a novel Saccharomyces cerevisiae sensor. *Biosensors and Bioelectronics*, 15(3–4), pp.211–219.

Lichty, J.J., Malecki, J.L., Agnew, H.D., Michelson-Horowitz, D.J. & Tan, S., 2005. Comparison of affinity tags for protein purification. *Protein Expression and Purification*, 41(1), pp.98–105.

Lippert, K. & Galinski, E., 1992. Enzyme stabilization be ectoine-type compatible solutes: protection against heating, freezing and drying. *Applied Microbiology and Biotechnology*, 37(1), pp.61–65.

Livrelli, V., Ike Whan Lee & Summers, A.O., 1993. In vivo DNA-protein interactions at the divergent mercury resistance (mer) promoters. I. Metalloregulatory protein MerR mutants. *Journal of Biological Chemistry*, 268(4), pp.2623–2631.

Lund, P.A. & Brown, N.L., 1989a. Regulation of transcription in Escherichia coli from the mer and merR promoters in the transposon Tn501. *Journal of Molecular Biology*, 205(2), pp.343–353.

Lund, P.A. & Brown, N.L., 1989b. Up-promoter mutations in the positively-regulated mer promoter of TNSOl. *Nucleic Acids Research*, 17(14), pp.5517–5528.

McAdams, H. & Arkin, A., 1997. Stochastic mechanisms in gene expression. *Proceedings of the National Academy of Sciences*, 94(3), pp.814–819.

Meech, S.R., 2009. Excited state reactions in fluorescent proteins. *Chemical Society Reviews*, 38(10), pp.2922–2934.

Merchant, B., 1998. Gold, the noble metal and the paradoxes of its toxicology. *Biologicals: journal of the International Association of Biological Standardization*, 26Merchant(1), pp.49–59.

Merulla, D., Buffi, N., Beggah, S., Truffer, F., Geiser, M., Renaud, P. & van der Meer, J.R., 2013. Bioreporters and biosensors for arsenic detection. Biotechnological solutions for a world-wide pollution problem. *Current Opinion in Biotechnology*, 24(3), pp.534–541.

Merulla, D., Hatzimanikatis, V. & Van der Meer, J.R., 2013. Tunable reporter signal production in feedback-uncoupled arsenic bioreporters. *Microbial Biotechnology*, 6(5), pp.503–514.

Milman, O., 2016. Millions exposed to dangerous lead levels in US drinking water, report finds. *The Guardian*.

Minton, A.P., 2000. Implications of macromolecular crowding for protein assembly. *Current Opinion in Structural Biology*, 10(1), pp.34–39.

Moore, C.M. & Helmann, J.D., 2005. Metal ion homeostasis in Bacillus subtilis. *Current opinion in microbiology*, 8(2), pp.188–195.

de Mora, K., Joshi, N., Balint, B.L., Ward, F.B., Elfick, A. & French, C.E., 2011. A pH-based biosensor for detection of arsenic in drinking water. *Analytical and bioanalytical chemistry*, 400(4), pp.1031–1039.

Morita, R.Y., Horikoshi, K. & Grant, W.D., 1999. Extremes of Biodiversity. *BioScience*, 49(3), p.245.

Naessens, M. & Tran-Minh, C., 1998. Whole-cell biosensor for direct determination of solvent vapours. *Biosensors and Bioelectronics*, 13(3–4), pp.341–346.

National Health and Medical Research Council, 2011. *Australian Drinking Water Guidelines 6*, Commonwealth of Australia.

Newberry, K.J. & Brennan, R.G., 2004. The structural mechanism of transcription activation by MerR family member MtaN. *J. Biol. Chem.*, 274, p.Published on the Web as Manuscript M400960200.

Nieman, T.A., Skoog, D.A. & Holler, F.J., 1998. *Principles of instrumental analysis*, Pacific Grove: Brooks/Cole.

Nies, D.H., 2003. Efflux-mediated heavy metal resistance in prokaryotes. *FEMS Microbiology Reviews*, 27(2–3), pp.313–339.

Nordberg, G.F., Fowler, B.A., Nordberg, M. & Friberg, L.T., 2007. *Handbook on the Toxicology of Metals*,

Nordblrom, D.K., 2002. Worldwide Occurrences of Arsenic i n Ground Water. *Science*, 296(June), pp.2143–2145.

Nriagu, J., 1996. A History of Global Metal Pollution. *Science*, 272(5259), p.223.

Olson, E., Fedinick, K.P., Egan, M. & Goffredi, L., 2016. *Whats in your water? Flint and beyond.*, Natural Resources Defense Council.

Osman, D. & Cavet, J.S., 2010. Bacterial metal-sensing proteins exemplified by ArsR-SmtB family repressors. *Natural product reports*, 27(5), pp.668–680.

Outten, C.E., Outten, F.W. & O'Halloran, T. V., 1999. DNA distortion mechanism for transcriptional activation by ZntR, a Zn(II)-responsive MerR homologue in Escherichia coli. *Journal of Biological Chemistry*, 274(53), pp.37517–37524.

Outten, F.W., Outten, C.E., Hale, J. & O'Halloran, T. V., 2000. Transcriptional activation of an Escherichia coli copper efflux regulon by the chromosomal MerR homologue, CueR. *Journal of Biological Chemistry*, 275(40), pp.31024–31029.

Paige, J.S., Wu, K.Y. & Jaffrey, S.R., 2011. RNA mimics of green fluorescent protein. *Science*, 333(6042), pp.642–6.

Pardee, K., Green, A.A., Ferrante, T., Cameron, D.E., Daleykeyser, A., Yin, P. & Collins, J.J., 2014. Paper-based synthetic gene networks. *Cell*, 159(4), pp.940–954.

Park, S.J., Wireman, J. & Summers, A.O., 1992. Genetic analysis of the Tn21 mer operator-promoter. *Journal of Bacteriology*, 174(7), pp.2160–2171.

Parkhill, J. & Brown, N.L., 1990. Site-specific insertion and deletion mutants in the mer promoter-operator region of Tn501; the nineteen base-pair spacer is essential for normal induction of the promoter by MerR. *Nucleic Acids Research*, 18(17), pp.5157–5162.

Patrick, L., 2006a. Lead toxicity, a review of the literature. Part 1: Exposure, evaluation, and treatment. *Alternative medicine review*, 11(1), pp.2–22.

Patrick, L., 2006b. Lead toxicity Part II: the role of free radical damageand the use of antioxidants in the pathology and treatment of leadtoxicity. *Alternative medicine review*, 11(2).

Pellinen, T., Huovinen, T. & Karp, M., 2004. A cell-free biosensor for the detection of transcriptional inducers using firefly luciferase as a reporter. *Analytical biochemistry*, 330(1), pp.52–57.

Philips, S.J., Canalizo-Hernandez, M., Yildirim, I., Schatz, G.C., Mondragón, A. & O'Halloran, T. V, 2015. Allosteric transcriptional regulation via changes in the overall topology of the core promoter. *Science*, 349(6250), pp.877–881.

Pizarro, F., Olivares, M., Uauy, R., Contreras, P., Rebelo, A. & Gidi, V., 1999. Acute gastrointestinal effects of graded levels of copper in drinking water. *Environmental Health Perspectives*, 107(2), pp.117–121.

Plum, L.M., Rink, L. & Hajo, H., 2010. The essential toxin: Impact of zinc on human health. *International Journal of Environmental Research and Public Health*, 7(4), pp.1342–1365.

Pohl, E., Haller, J.C., Mijovilovich, A., Meyer-Klaucke, W., Garman, E. & Vasil, M.L., 2003. Architecture of a protein central to iron homeostasis:crystal Structure and Spectroscopic analysis of the ferric uptate regulator. *Molecular Microbiology*, 47(4), p.903.

Ponchon, L. & Dardel, F., 2007. Recombinant RNA technology: the tRNA scaffold. *Nature methods*, 4(7), pp.571–576.

Pooley, D.T., Larsson, J., Jones, G., Rayner-Brandes, M.H., Lloyd, D., Gibson, C. & Stewart, W.R., 2004. Continuous culture of photobacterium. *Biosensors and Bioelectronics*, 19(11), pp.1457–1463.

Pothoulakis, G., Ceroni, F., Reeve, B. & Ellis, T., 2013. The Spinach RNA aptamer as a characterisation tool for synthetic biology. *ACS Synthetic Biology*, 3(3), pp.182–187.

Promega Corporation, 2015a. E. coli S30 Extract System for Circular DNA - Technical Bulletin, TB092. , p.17.

Promega Corporation, 2015b. pGEM®- T and pGEM®- T Easy Vector Systems. *Technical Manual*, (TM042), pp.1–28.

Rabinowitz, M.B., 1991. Toxicokinetics of bone lead. *Environmental Health Perspectives*, 91, pp.33–37.

Ralston, D.M. & O'Halloran, T. V., 1990. Ultrasensitivity and heavy-metal selectivity of the allosterically modulated MerR transcription complex. *Proceedings of the National Academy of Sciences of the United States of America*, 87(10), pp.3846–3850.

Ramos, J.A., Bermejo, E., Zapardiel, A., Pérez, J.A. & Hernández, L., 1993. Direct determination of lead by bioaccumulation at a moss-modified carbon paste electrode. *Analytica Chimica Acta*, 273(1), pp.219–227.

Rawson, D., Willmer, A. & Turner, A., 1989. Whole-cell biosensors for environmental monitoring. *Biosensors*, 4(5), pp.299–311.

Rice, E.W., Baird, R.B., Eaton, A.D. & Clesceri, L.S. eds., 2012. *Standard Methods for the Examination of Water and Wastewater* 22nd ed., American Public Health Association, American Water Works Association, Water Environment Federation.

Richins, R.D., Kaneva, I., Mulchandani, A. & Chen, W., 1997. Biodegradation of organophosphorus pesticides by surface-expressed organophosphorus hydrolase. *Nature biotechnology*, 15(10), pp.984–7.

Riedel, K., Kunze, G. & König, A., 2002. Microbial sensors on a respiratory basis for wastewater monitoring. *Advances in biochemical engineering/biotechnology*, 75, pp.81–118.

Rochette, E.A., Bostick, B.C., Li, G. & Fendorf, S., 2000. Kinetics of Arsenate Reduction by Dissolved Sulfide. *Environmental Science & Technology*, 34(22), pp.4714–4720.

Rosen, B. & Borbolla, M., 1984. A plasmid-encoded arsenite pump produces arsenite resistance in Escherichia coli. *Biochemical and biophysical research communications*, 124(3), pp.760–765.

Ruan, X., Bhattacharjee, H. & Rosen, B.P., 2006. Cys-113 and Cys-422 form a high affinity metalloid binding site in the ArsA ATPase. *Journal of Biological Chemistry*, 281(15), pp.9925–9934.

Salis, H.M., Mirsky, E.A. & Voigt, C.A., 2009. Automated design of synthetic ribosome binding sites to control protein expression. *Nature biotechnology*, 27(10), pp.946–50.

Sambrook, J., Fritsch, E. & Maniatis, T., 1989. *Molecular cloning*, New York: Cold Spring Harbor.

Sanger, F., Nicklen, S. & Coulson, A.R., 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, 74(12), pp.5463–7.

Sarkar, B. ed., 2002. *Heavy metals in the environment* 1st ed., New York: Marcel Dekker.

Satterlee, H. & Blodgett, G., 1944. Ultramicrodetermination of Arsenic by Gutzeit Spot-Filtration under Vacuum. A Rapid Technique Employing Photometric Calibration and Permanent Photographic. *Industrial & Engineering Chemistry Analytical Edition*, 16(6), pp.400–407.

Shaner, N.C., Steinbach, P. a & Tsien, R.Y., 2005. A guide to choosing fluorescent proteins. *Nature methods*, 2(12), pp.905–909.

Sheppard, M.I., 1993. Heavy Metals in the Environment L. K. Wang et al., eds. *Journal of Environment Quality*, 22(1), p.213.

Shi, W., Dong, J., Scott, R. a, Ksenzenko, M.Y. & Rosen, B.P., 1996. The role of arsenic-thiol interactions in metalloregulation of the ars operon. *The Journal of biological chemistry*, 271(16), pp.9291–9297.

Shi, W., Wu, J. & Rosen, B.P., 1994. Identification of a putative metal binding site in a new family of metalloregulatory proteins. *Journal of Biological Chemistry*, 269(31), pp.19826–19829.

Shin, J. & Noireaux, V., 2010. Efficient cell-free expression with the endogenous E. Coli RNA polymerase and sigma factor 70. *Journal of biological engineering*, 4(1), p.8.

Shrestha, P., Holland, T.M. & Bundy, B.C., 2012. Streamlined extract preparation for Escherichia coli-based cell-free protein synthesis by sonication or bead vortex mixing. *BioTechniques*, 53(3), pp.163–174.

Siddiki, M.S.R., Kawakami, Y., Ueda, S. & Maeda, I., 2011. Solid phase biosensors for arsenic or cadmium composed of A trans factor and cis element complex. *Sensors*, 11(11), pp.10063–73.

Siddiki, M.S.R., Ueda, S. & Maeda, I., 2012. Fluorescent bioassays for toxic metals in milk and yoghurt. *BMC biotechnology*, 12(1), p.76.

Siegel, E., 2002. *Environmental Geoehemistry of Potentially Toxie Metals*,

Silver, S., Budd, K. & Leahy, K., 1981. Inducible plasmid-determined resistance to arsenate, arsenite, and antimony (III) in escherichia coli and Staphylococcus aureus. *Journal of bacteriology*, 146(3), p.983.

Simons, R.W., Houman, F. & Kleckner, N., 1987. Improved single and multicopy lac-based cloning vectors for protein and operon fusions. *Gene*, 53(1), pp.85–96.

Simpson, M. & Sayler, G., 2001. Whole-cell biocomputing. *TRENDS in Biotechnology*, 19(8), pp.317–323.

Smith, A.H., Lingas, E.O. & Rahman, M., 2000. Contamination of drinking-water by arsenic in Bangladesh: a public health emergency. *Bulletin of the World Health Organisation*, 78(9), pp.1093–1103.

Snyder, L.J., 1947. Improved Dithizone Method for Determination of Lead. *Analytical Chemistry*, 19(9), pp.684–687.

Song, W., Strack, R.L., Svensen, N. & Jaffrey, S.R., 2014. Plug-and-play fluorophores extend the spectral properties of Spinach. *Journal of the American Chemical Society*, 136(4), pp.1198–201.

Spector, J.T., Navas-Acien, A., Fadrowski, J., Guallar, E., Jaar, B. & Weaver, V.M., 2011. Associations of blood lead with estimated glomerular filtration rate using MDRD, CKD-EPI and serum cystatin C-based equations. *Nephrology Dialysis Transplantation*, 26(9), pp.2786–2792.

Stepanenko, O., Stepanenko, O., Shcherbakova, D., Kuznetsova, I., Turoverov, K. & Verkhusha, V., 2011. Modern fluorescent proteins: from chromophore formation to novel intracellular applications. *BioTechniques*, 51(5), p.313–4, 316, 318 passim.

Stothard, P., 2000. The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *BioTechniques*, 28(6), pp.1102, 1104.

Stoyanov, J. V. & Brown, N.L., 2003. The Escherichia coli Copper-responsivecopA Promoter Is Activated by Gold. *Journal of Biological Chemistry*, 278(3), pp.1407–1410.

Stoyanov, J. V, Hobman, J.L. & Brown, N.L., 2001. CueR (YbbI) of Escherichia coli is a MerR family regulator controlling expression of the copper exporter CopA. *Molecular Microbiology*, 39(2), pp.502–511.

Stoyanov, J. V, Magnani, D. & Solioz, M., 2003. Measurement of cytoplasmic copper, silver, and gold with a lux biosensor shows copper and silver, but not gold, efflux by the CopA ATPase of Escherichia coli. *FEBS Letters*, 546(2–3), pp.391–394.

Strack, R.L., Disney, M.D. & Jaffrey, S.R., 2013. A superfolding Spinach2 reveals the dynamic nature of trinucleotide repeat-containing RNA. *Nature methods*, 10(12), pp.1219–24.

Struss, A., Pasini, P., Ensor, C.M., Raut, N. & Daunert, S., 2010. Paper strip whole cell biosensors: A portable test for the semiquantitative detection of bacterial quorum signaling molecules. *Analytical Chemistry*, 82(11), pp.4457–4463.

Studier, F.W., 2005. Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification*, 41(1), pp.207–234.

Su, L., Jia, W., Hou, C. & Lei, Y., 2011. Microbial biosensors: A review. *Biosensors and Bioelectronics*, 26(5), pp.1788–1799.

Sun, Y., Perch-Nielsen, I., Dufva, M., Sabourin, D., Bang, D.D., Høgberg, J. & Wolff, A., 2012. Direct immobilization of DNA probes on non-modified plastics by UV irradiation and integration in microfluidic devices for rapid bioassay. *Analytical and Bioanalytical Chemistry*, 402(2), pp.741–748.

Sun, Z.Z., Hayes, C.A., Shin, J., Caschera, F., Murray, R.M. & Noireaux, V., 2013. Protocols for implementing an Escherichia coli based TX-TL cell-free expression system for synthetic biology. *Journal of visualized experiments: JoVE*, 50762(79), p.e50762.

Sweeney, T., 2007. Auto-induction for over expression in E. coli. *Centre for Structural Biology Techniques Workshop on Cloning and Expression*, pp.1–18. Available at: http://www3.imperial.ac.uk/pls/portallive/docs/1/15699698.PPT.

Takeuchi, T., Morikawa, N., Matsumoto, H. & Shiraishi, Y., 1962. A pathological study of Minamata disease in Japan. *Acta Neuropathologica*, 2(1), pp.40–57.

Taniguchi, M., Siddiki, M.S.R., Ueda, S. & Maeda, I., 2014. Mercury (II) sensor based on monitoring dissociation rate of the trans-acting factor MerR from cis-element by surface plasmon resonance. *Biosensors & bioelectronics*, (Ii), pp.1–6.

Tanner, M.S., 1998. Role of copper in indian childhood cirrhosis. In *American Journal of Clinical Nutrition*. p. 1074S–1081S.

Tartof, K. & Hobbs, C., 1987. Improved media for growing plasmid and cosmid clones. *Focus*, 9(2), p.10.

Taylor, M.T., Belgrader, P., Furman, B.J., Pourahmadi, F., Kovacs, G.T.A. & Northrup, M.A., 2001. Lysing bacterial spores by sonication through a flexible interface in a microfluidic system. *Analytical Chemistry*, 73(3), pp.492–496.

Thevenon, F., Graham, N.D., Chiaradia, M., Arpagaus, P., Wildi, W. & Poté, J., 2011. Local to regional scale industrial heavy metal pollution recorded in sediments of large freshwater lakes in central Europe (lakes Geneva and Lucerne) over the last centuries. *Science of the Total Environment*, 412–413, pp.239–247.

Tian, J.Q. and J. & Quan, J., 2009. Circular Polymerase Extension Cloning of Complex Gene Libraries and Pathways P. L. Ho, ed. *PLoS One*, 4(7), p.e6441.

Trinh, R., Gurbaxani, B., Morrison, S.L. & Seyfzadeh, M., 2004. Optimization of codon pair use within the (GGGGS)3 linker sequence results in enhanced

protein expression. *Molecular Immunology*, 40(10), pp.717–722.

Trubitsyna, M., Michlewski, G., Cai, Y., Elfick, A. & French, C.E., 2014. PaperClip: rapid multi-part DNA assembly from existing libraries. *Nucleic acids research*, 42(20), p.e154.

Turner, A., Karube, I. & Wilson, G.S., 1987. Biosensors: Fundamentals and Applications A. Turner, I. Karube, & G. S. Wilson, eds. , p.770.

Turner, A.P.F., 2000. Biosensors--Sense and Sensitivity. *Science*, 290(5495), pp.1315–1317.

Utschig, L., Bryson, J. & Halloran, T., 1995. Mercury-199 NMR of the metal receptor site in MerR and its protein-DNA complex. *Science*.

Verma, N. & Singh, M., 2005. Biosensors for heavy metals. *BioMetals*, 18(2), pp.121–129.

Vossen, K.M. & Fried, M.G., 1997. Sequestration stabilizes lac repressor-DNA complexes during gel electrophoresis. *Analytical biochemistry*, 245(1), pp.85–92.

Wang, J., 2001. Glucose Biosensors: 40 Years of Advances and Challenges. *Electroanalysis*, 13(12), pp.983–988.

Wang, X.C., Wilson, S.C. & Hammond, M.C., 2016. Next-generation RNA-based fluorescent biosensors enable anaerobic detection of cyclic di-GMP. *Nucleic Acids Research*, 44(17), p.gkw580.

Watanabe, S., Kita, A., Kobayashi, K. & Miki, K., 2008. Crystal structure of the [2Fe-2S] oxidative-stress sensor SoxR bound to DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 105(11), pp.4121–4126.

Wegner, S. V., Okesli, A., Chen, P. & He, C., 2007. Design of an emission ratiometric biosensor from MerR family proteins: A sensitive and selective sensor for Hg2+. *Journal of the American Chemical Society*, 129(12), pp.3474–3475.

World Health Organization, Arsenic. Available at: http://www.who.int/mediacentre/factsheets/fs372/en/index.html [Accessed August 8, 2013].

World Health Organization, 1996. *Guidelines for drinking-water quality*, Geneva.

World Health Organization, 2013. Water-related diseases. Available at: http://www.who.int/water_sanitation_health/diseases/arsenicosis/en/ [Accessed August 8, 2013].

World Health Organization, Aggett, P.J., Beaton, G.H., Dreosti, I., Goyer, R.A., Hetzel, B.S., Iyengar, G. V, Kostial, K., Krishnamachari, K.A.V.R., Levander, O.A., Mertz, W., Mills, C.F., Nielsen, W.H., Parr, R.M., Rish, M.A. & Sandstroem, A.B., 1996. *Trace elements in human nutrition and health*, Geneva: World Health Organization.

Wu, J. & Rosen, B.P., 1993. Metalloregulated expression of the ars operon. *Journal of Biological Chemistry*, 268(1), pp.52–58.

Wu, J. & Rosen, B.P., 1993. The arsD gene encodes a second trans-acting regulatory protein of the plasmid-encoded arsenical resistance operon. *Molecular microbiology*, 8(3), pp.615–23.

Wurch, T., Lestienne, F. & Pauwels, P.J., 1998. A modified overlap extension PCR method to create chimeric genes in the absence of restriction enzymes. *Biotechnology Techniques*, 12(9), pp.653–657.

Xu, C., Shi, W. & Rosen, B.P., 1996. The chromosomal arsR gene of Escherichia coli encodes a trans-acting metalloregulatory protein. *The Journal of biological chemistry*, 271(5), pp.2427–32.

Yagi, K., 2007. Applications of whole-cell bacterial sensors in biotechnology and environmental science. *Applied Microbiology and Biotechnology*, 73(6), pp.1251–1258.

Yamamoto, S. & Sano, Y., 1992. Drying of enzymes: enzyme retention during drying of a single droplet. *Chemical Engineering Science*, 47(1), pp.177–183.

Yamasaki, A., Cunha, M.Â.S.D.A., Oliveira, J.A.B.P., Duarte, A.C. & Gomes, M.T.S.R., 2004. Assessment of copper toxicity using an acoustic wave sensor. *Biosensors and Bioelectronics*, 19(10), pp.1203–1208.

Yáñez-Sedeño, P., Agüí, L., Villalonga, R. & Pingarrón, J.M., 2014. Biosensors in forensic analysis. A review. *Analytica Chimica Acta*, 823, pp.1–19.

Yoo, E.-H. & Lee, S.-Y., 2010. Glucose Biosensors: An Overview of Use in Clinical Practice. *Sensors*, 10(5), pp.4558–4576.

Zhang, J., Sonnenschein, N., Pihl, T.P.B., Pedersen, K.R., Jensen, M.K. & Keasling, J.D., 2016. Engineering an NADPH/NADP+ redox biosensor in yeast. *ACS Synthetic Biology*, p.acssynbio.6b00135.

Zubay, G., 1973. In vitro synthesis of protein in microbial systems. *Annual review of genetics*, 7, pp.267–87.